THE NON-SLAVIC LANGUAGES OF THE USSR

Linguistic Studies
New Series

Edited by
HOWARD I. ARONSON
University of Chicago

Chicago Linguistic Society
University of Chicago
1992

THE NON-SLAVIC LANGUAGES OF THE USSR

Linguistic Studies

Second Series

Edited by
HOWARD I. ARONSON
University of Chicago

Chicago Linguistic Society
University of Chicago
1992

Published 1992 by:
THE CHICAGO LINGUISTIC SOCIETY
University of Chicago
1050 East 59th Street
Chicago, Illinois 60637

Each contribution copyright © 1992 by its author. All rights reserved.

The editor and the publisher assume no responsibility for statements of fact or opinion by the authors.

ISBN: ISBN 0-914203-41-X

Printed in the United States of America.

PREFACE

Most of the papers in this volume were originally presented in abbreviated and, often, in preliminary form at the Sixth International Conference on the Non-Slavic Languages of the USSR, held at the University of Chicago in May, 1989. There are a number of people who have helped greatly in the processes leading to the publication of this volume and I would like to express my gratitude to them: to my colleague, Bill J. Darden, co-organizer of the Conference; to Philip Gossett, Dean of the Division of the Humanities of the University of Chicago, for helping to make the publication of the volume possible; to the officers of the Chicago Linguistic Society for agreeing to have CLS serve as the publisher of the volume; to David Testen, for his generous aid in the organization of the Conference and the preparation of the volume; and to all the contributors to the volume. A number of the papers were subjected to outside evaluation and I would like to also thank those who served as referees.

When the participants in this conference met, the Union of Soviet Socialist Republics still existed. Now there are fifteen new, independent republics, all of which are multi-ethnic and multilingual, and most of which are little known and little studied in the West. We hope that these linguistic studies will serve not only to make the study of the peoples, languages, and cultures of this area more accessible to all but also to expand our understanding of the structure of language and languages.

Howard I. Aronson
Chicago, July 1992

TABLE OF CONTENTS

Howard I. Aronson. On the marking of O_3 in Georgian 1

John Colarusso. Phyletic links between proto-Indo-European and proto-Northwest Caucasian 19

Bernard Comrie. Focus in Yukaghir (Tundra dialect) 55

Bill J. Darden. On the syntax of Lithuanian participles 71

Michael Dobrovolsky. Towards a lexical phonology of Chuvash: inflection as derivation 99

Victor A. Friedman. Lak substantival declension: 40 cases or 50?. 113

Alice C. Harris. The particle *-a* in Udi 133

Maria Polinsky. CrimeanTatar and Krymchak: classification and description 157

Wolfgang Schulze-Fürhof. How can class markers petrify? Towards a functional diachrony of morphological subsystems in the East Caucasian languages 189

David Testen. Ibn Rusta and Bakrī on the Ossetians and the Apkhaz 235

Kevin Tuite. The category of number in Common Kartvelian....... 245

Steven Young. Indo-European length and Baltic tone 285

ON THE MARKING OF O_3 IN GEORGIAN

Howard I. Aronson
University of Chicago

The purpose of this paper is both simple and quite limited: to determine whether in contemporary Georgian we can speak about the existence of a **morphological** indication in the verb form of the presence of a third person direct or indirect object. Most grammars and textbooks of Georgian treat the verb as polypersonal, marking the person of subject, direct object (DO), and indirect object (IdO).[1] But this polypersonalism is complicated by what Gamqrelije (1981) has called the prefixal monopersonalism of Georgian: all marking of person[2] in Georgian is prefixal, with the exception of the suffixal 3rd person subject markers.[3] However, the marking of person in the Georgian verb is not agglutinative because of the rule that in general allows only one personal prefix in a verb form. Determination of the persons marked by a verb form is, consequently, a fairly complex process requiring consideration of what is and what is *not* marked in the verb form. As an example, consider the following forms:

> *gxedav* 'I see you.' The presence of the prefix *g-* indicates a 2nd person object. The **absence** of a suffix indicates that the subject marker is not 3rd person. Given a rule against coreferential 1st and 2nd persons within a given verb form, the subject must be first person, despite the fact that there is no overt marker *v-* in the verb form (cf. *vxedav* 'I see him').
>
> *mxedav* 'you see me.' The prefix *m-* indicates the presence of a first person object. The absence of a suffix indicates that the subject marker is not 3rd person and, given the rule against coreference, the subject must be 2nd person, which normally does not have a formal marker.[4]

In what follows we shall discuss the zero marking of DO_3, the so-called *h*-series indirect object markers, and the preradical vowels, all of which have been viewed as giving positive indication of the presence of an object.

DO_3 = zero. Traditionally, in transitive (I. conjugation) verbs, in the present/future and aorist series, the third person direct object, both singular and plural, is viewed as having a zero-marker, unlike the first and second persons:

1.

	SINGULAR	PLURAL
1.	m-	gv-
2.	g-	g-...-t
3.	Ø-	

Example:

1sg.	m-xedavs	he sees me
1pl.	gv-xedavs	he sees us
2sg.	g-xedavs	he sees you
2pl.	g-xedavt	he sees you all
3.	Ø-xedavs	he sees her/them

According to this analysis, the absence of an object marker in the third person is distinctive, a positive indication of the presence of a third person direct object; only the category of number is neutralized. But, there are instances when it is clear that a transitive I conjugation verb can occur without a direct object; e.g.,

2.a. Ğamit tagvebi ćudad xedaven. 'Mice see poorly at night.' (KEGL)
2.b. Arsaķije uķet xaţavs vidre ašenebs. 'Arsaķije paints better than he builds.' (KEGL)

Compare also the following:

3.a. **Daaxasiata** masçavlebelma moçape. 'The teacher characterizes the student [DO].'
3.b. Am moçapes ķargi čveva **axasiatebs**. 'This student [DAT] is characterized by good habits.' (J̌orbenaje 1983:82.)

In the first two examples there is no direct object, although the verbs, *xedaven, xaṭavs, ašenebs* are transitive. In the last pair, the (a.) sentence can only occur with a direct object, while the (b.) sentence cannot ever take a direct object. Similar examples include:

4.a. Gaiozma **daabraxuna ḳarebi.** 'Gaioz knocked on the **door** [DO].' (KEGL)
4.b. Giorgi mankanis ḳabinaze **daabraxunebs**... 'Giorgi knocks on the cabin of the truck.' (J̌.m. 136)

5.a. Ekimis **bina ačvena.** 'He pointed out the doctor's **apartment** [DO].' (KEGL)
5.b. Xelit **ačvenebs** ǰarisḳacebze. 'With his hand he indicates the soldiers.' (J̌.m. 120)

6.a. **Samotxesac** akedamve eg **anatebs.** 'This illuminates even **paradise** [DO] from here.' (Suxišvili 1979:58)
6.b. Tkven icit, rom kveqanazed mašin upro tbila, roca mze **anatebs.** 'You know that it's warmer on earth when the sun shines.' (Suxišvili 1979:58)

In the (a.) sentences, the verb has a direct object (indicated in boldface); in the (b.) sentences, there is no direct object. It is clearly meaningless to posit an opposition of the type *da-Ø-a-braxun-a ḳarebi // da-a-braxun-eb-s; Ø-a-čven-a bina // Ø-a-čven-eb-s, Ø-a-nat-eb-s samotxesa // a-nat-eb-s*. Therefore, we cannot speak of a zero morpheme marking DO$_3$. Such a solution would result in a contrast between a zero marker Ø- which never has a non-zero variant and the absence of this zero marker. In such a situation the zero posited above cannot be regarded as a distinctive, meaningful absence of marking. Simply put, a form such as *xedavs* gives *no* formal information as to the presence or absence of a direct object, while forms such as *mxedavs, gvxedavs, gxedavs, gxedavt* give unambiguous information as to the presence of an object.[5] A Georgian verb form such as *ċers* 'writes' gives no more and no less information about the presence or absence of a direct object than does the corresponding Russian verb form *pišet* 'writes.'

In light of the above, it is necessary to restate the rules for the marking of direct objects in Georgian: Georgian verbs can indicate the presence of a *personal* direct object, where personal is defined as the first and second persons.[6] The paradigm is thus:

7.	MARKING FOR D.O. OBLIGATORILY TRANSITIVE		NO MARKING FOR D.O. TRANSITIVE OR INTRANSITIVE
	m-xedavs	gv-xedavs	xedav-s
	g-xedavs	g-xedavt	

Further evidence of the lack of marking of DO_3 can be seen in those verbs which are labile in the third person and can be used both transitively and intransitively, e.g., *miic̣evs:* (a.) 's.o. will pull sthg. to oneself'; (b.) 's.o. is moving.' Clearly, it is not reasonable to posit a "hocus-pocus" opposition between a *mi-Ø-i-c̣ev-s* and a *mi- -i-c̣ev-s.*

A parallel situation is found in the perfect series of I. and III. conjugation verbs, where an overt marker *(-s* or *-a* in the present perfect, *-a* in the pluperfect), usually treated as DO_3, similarly does not necessarily indicate the presence of a direct object in the nominative as, for example, in III. conjugation verbs.

8.a. PERFECT	PLUPERFECT	PRESENT/FUTURE	GDW[7]	
udu${\breve{g}}$nia	edu${\breve{g}}$na	du${\breve{g}}$s	MV	boil
ulaṗaraḳnia	elaṗaraḳna	laṗaraḳobs	MV	speak
gauṭevia	gaeṭeva	gauṭevs	T^3	hurry; curse s.o.

Cf. also the following, where there is no nominative direct object despite the presence of the formant *-a:*

8.b. Monadiris mosvla aravis gahḳvirvebia, arcvis-ḳi **šeuxednia** [= šeuxedavs] imistvis. 'The hunter's arrival didn't astonish anyone; no one even paid-attention to it.' (Važa in KEGL.)

It is worth noting that the lack of marking for a third person direct object in transitive verbs finds a partial parallel in the the S_3 marker, which does not necessarily indicate the presence of a subject noun phrase; in impersonal constructions the S_3 does not indicate such a presence. Examples include: *c̣vim-s* 'it is raining,' *tov-s* ' it is snowing,' *civ-a* 'it is cold,' etc. (See Benveniste 1966: 230.)

It may be objected that in many (though not all) instances the absence of a direct object with transitive verbs is possible only in present[8] series screeves; note the possibility of *xaṭavs* 'he paints' as opposed to **daxaṭa* 'he painted' without direct object. But as in German, Russian, and other languages, Georgian transitive verbs which allow intransitive constructions (without change in the semantic function of the grammatical subject) are generally imperfective (for Georgian, usually unprefixed) and are capable of an activity reading in terms of Vendler's categories (Vendler 1967). Prefixed forms of the verb become telic and thus generally incapable of an activity reading. The obligatoriness of a direct object is, then, a function of lexical or grammatical aspect and not necessarily of (inherent) transitivity.

Indirect object markers. If there is no marker of a third person *direct* object in Georgian, the question can be raised as to the status of third person *indirect* object markers. The grammars and textbooks generally give the following indirect object marker series:

9. *H-* SERIES		*A-* SERIES[9]	
1. m-	gv-	1. ma-	gva-
2. g-	g-...-t	2. ga-	ga-...-t
3. h-/s-/Ø-		3. a-	

U- SERIES		*E-* SERIES	
1. mi-	gvi-	1. me-	gve-
2. gi-	gi-...-t	2. ge-	ge-...-t
3. u-		3. e-	

***H-* series.** The occurrence of the various O_3 forms of the *h*-series in normative Georgian is phonologically conditioned: *s-* occurs before /d, t, ṭ; j, c, c̣; ǰ, č, č̣/, *h-* before /g, k, ḳ; q; p̣/ and facultatively (though rarely) before all other consonants; *Ø*- is obligatory before all vowels and facultative (though the rule) before consonants other than those taking *s*- and those "obligatorily" taking *h*-. Such is the normative rule. But in actual usage, we find two deviations: (a.) the absence of *h*-, *s*- in forms containing a third person indirect object, and (b.) the occurrence of *h*-, *s*- in forms without a third person indirect object.

Examples of (a.):

10. ...oci c̣lis unaxavi megobari davḳarge, magram saplavši ar **čavqolivar** [= čavhqolivar]! 'I lost a friend I hadn't seen for 20 years, but I didn't follow him (IdO) to the grave.' (M.ğ.)

11. Sṭipendia pulis am gansacviprebel ḳaṭegorias eḳutvnis, romelic cidan čamovardnils **gavs** [= **h**gavs]. 'A stipend belongs to that remarkable category of money that resembles something fallen down from heaven.' (M.ğ.)

Examples of (b.):

12. Is ṗiesa, romelic am sami c̣lis c̣inat **dasc̣era** [= dac̣era]... 'The play which he had written some three years ago...' (Sicili)

13. Man ṭirili daic̣qo da momaḳvdavis otaxi **dasṭova** [= daṭova]. 'He began to cry and left the dying person's room.' (Sicili)

14. Romeliğac saplavze prtebč̣amoqrilma sicilma šeiprtxiala, ḳamara **šehḳra** [= šeḳra], sazogadoebaši šeič̣ra da ğimili mohpina damc̣uxrebul saxeebs. 'At a certain grave a desperate laughter began to flutter, surrounded the vault, tore into the crowd and covered the grieving faces with a smile.' (Sicili)

The third variant of the IdO$_3$ marker is *Ø*-. Arguments similar to those given above, arguing against the existence of a DO$_3$ marker, can be given against the existence of an IdO$_3$ = *Ø*-. Compare, for example, the following:

15. [Teo] **dabanes,** ... šeamšrales. ‘They washed Teo [DO], ...they wiped him dry.’ (KEGL)

16. [Švils] cxvir-ṗiri... **dabana** da moalamaza. ‘He washed the child’s [IdO] face [DO] and made him look better.’ [KEGL]

In the first sentence there is a direct object, but no indirect object. In the second, there is both a direct object and an indirect object. Yet, there is no indication in the verb form of the presence or absence of an indirect object. We can, again, only conclude that a zero prefix indicating an indirect object cannot be posited.

Thus, neither the presence nor the absence of *h-, s-* gives information as to the presence or absence of a third person indirect object. The first and second person markers do, however, indicate the presence of a personal **object,** although, since these forms are homonymous with the corresponding first and second person direct object markers, they cannot give an unambiguous *formal* indication of whether the object is direct or indirect.[10] We conclude, again, that there is no marking of third person objects in the verb form.

Preradical vowels. The preradical vowels *a-, e-,* and *u-* are said to mark, in certain forms, the presence of an object, direct *(a-)* or indirect *(a-, e-, u-).* Some scholars (e.g., Tschenkéli in GDW) posit homonymous forms, e.g., *a-*$_1$marking ‘neutral version’ and *a-*$_2$, marking the ‘superessive.’ Similarly one could posit, e.g., three homonymous forms *e-,* one marking an indirect object in relative II. verbs *(daeç̣ereba),* one marking the pluperfect of I. and III. conjugation verbs *(daeç̣era),* and one marking the future and aorist series of IV. conjugation verbs *(eç̣ereba).* Our approach, however, is to avoid recourse to homonymy, but rather to rely for our analysis on whatever information is available on the surface level. Thus, since on the surface there is no formal way of differentiating between the relative II. conjugation aorist *gaegzavna* ‘it was sent to him‘ and the I. conjugation pluperfect *gaegzavna* ‘[that] he had sent it,’ or between the II. conjugation present form *eç̣ereba* ‘it is written on sthg.’ and the future of the passive of state *eç̣ereba* ‘it will be written,’ our analysis will not posit a different morpheme *e-* in the two forms. We thus operate with only one *a-*, one *e-*, and one *u-*

***A*-series.** The *a-* series indirect object markers (GDW: T^4 and T^5) are homonymous with sequences of direct object marker and the preradical vowel *a-*, which usually has transitivizing function. Consequently, again, the presence of the sequences *ma-*, *ga-*, *gva-* indicates the presence of an object, but with I. conjugation verbs it is generally impossible to predict on formal grounds whether the object is direct or indirect. The third person *a-*, however, can occur with verbs correlated with a direct object, an indirect object, or with the absence of any object. Examples:

17.a. Sabç̌ota armiam dağuṗvas **gadaarčina** mteli Evroṗis xalxebi. 'The Soviet army saved the peoples [DO] of all Europe from destruction.' (KEGL)
17.b. Gulçasuli ṗartizanma moiṗoves da **gadamarčines.** 'Let the partisan find the man who fainted and let him save me [DO].' (KEGL)

In the above, example (a.) has an indirect object, while example (b.) has none.

18.a. Viğacam **daaxvela.** 'Someone coughed.' (KEGL)
18.b. Miirtvi, švilo, miirtvi. Bevri-ķi ara sç̌amo, torem qels dagçvavs da **dagaxvelebs.** 'Take something to eat, child, take something to eat. But don't eat [too] much or else it will burn your throat and you [DAT] will cough.' (KEGL)

In the above, there is no actant correlated with the preradical vowel *a-* in (a.), while in (b.) there is a 2nd person dative (marked by the prefix *g-).*[11] A similar situation is found when one contrasts the two passives of state below:

19. antia	sthg is burning (no object)
açeria	it is written on sthg. (dative object)

Again, there is a marking of personal objects ("personal" understood as 1st and 2nd person) but no marking of non-personal (third person)[12] objects.

***U-* series markers.** The personal *u-* series markers, *mi-, gvi-, gi-,* are the most distinctive of the indirect object markers discussed up to now. Nevertheless, they can be homonymous with the sequence of "direct object marker" plus the preradical vowel *i-* (GDW: T^2) as in such forms as: *Ekimi mimiğebs* 'the doctor will receive me,' where the *m-* marks a direct object and not an indirect object. The O_3 form, *u-*, too, is the most distinctive of the third person indirect object markers discussed up to now: unlike *h-, s-,* it is not facultative in normal usage and it lacks the ambiguity of *a-*. In the overwhelming majority of its occurrences, *u- does* indicate the presence of a third person indirect object. It is the marker of third person indirect object *par excellence* for I. and III. conjugation verbs and for II. conjugation verbs in *-d-*.

And yet, there are some instances when *u-* occurs without the marking of an indirect object. It, too, then, cannot be viewed as having the marking of a dative actant as its *invariant* meaning.[13] Examples of such usage are:

20. Me ţipi makvs, amas ç̌leki! mi**u**tita čemze Artavazam. 'I have typhus and this one has tuberculosis,' Artavaz pointed to me. (M.ğ.)

21. Es saati čemi mxridan ğamis ṗirvel saats **u**čvenebs. 'This clock from my side shows one a.m.' (M.ğ.)

22. Mačabelma ǰer mxariulad čonguri da**u**ḳra da mere saocari sisçrapit šeuţia sṗarselebs. 'Mačabeli first played the čonguri merrily and then with astonishing rapidity attacked the Persians. (KEGL)

The presence of *m-, gv-,* or *g-* in a verb form is an unambiguous indication of the presence of an object; the nature of that object, i.e.,whether it is direct or indirect is often formally unpredictable. The only prefix that can be viewed as a clear marker of a third person indirect object is *u*-, even though the predictability is not absolute, and, as we have seen, there are exceptional instances when the *u*- is not correlated with a DAT noun phrase.

***E*- series markers.** Unlike the indirect object markers discussed above, which occur readily with I. conjugation verbs (and, with respect to *h*-and *u*- series markers, also with III. conjugation verbs) in the present, future, and aorist, the *e*- series object markers are extremely rare with I. and III. conjugation verbs, e.g., *momeci*[14] 'you gave it to me' and *vejebdi* 'I was looking for it.'

E- series markers most commonly occur as the relative counterparts of absolute II. conjugation verbs in *i*-.[15] They are also markers of the grammatical subject in the pluperfect of I. and III. conjugation verbs and occur in the future and aorist series of passives of state and other IV. conjugation verbs.[16] We shall restrict our discussion to those instances where the *e*-marker occurs without a third person dative object.

***Eniani*-II. conjugation.** Absolute II. conjugation verbs in *i*- can become relative, i.e., add a dative object, with the change of the preradical vowel *i*- to *e*-. Examples:

23. **Igzavnebodnen** šiḳriḳebi Sakartvelos qvela ḳutxisḳen. 'Couriers were being sent to every corner of Georgia.' (KEGL)

24. Gamovida da xelis momçert [IdO] **egzavnebat** Didi sabçota enciḳlopediis ṭomebi. 'It came out and the volumes of the *Great Soviet Encyclopedia* were sent to the subscribers.' (KEGL)

But there are many instances of II. conjugation verbs with the O_3 preradical vowel *e*- that are not relative, but absolute, i.e., they do not occur with a dative object.[17] Examples of these include:

25. gaešureba hurry off:

Qvelani cixis ḳarebisḳen **gaešurnen**. 'All hurried off toward the door of the fortress.' (KEGL)

gaemgzavreba travel off
Magdana kalakisḳen gaemgzavra. 'He traveled off toward the town of Magdana.' (KEGL)

moekceva find oneself, end up, reach (a place)
Sakartvelo ... or mojalades šoris moekca. 'Georgia ended up between two oppresors.' (KEGL)

One and the same *eniani*-form can be used both with an indirect object and without:

26.a. Ḳodala vis rad undoda, ṭqavs aravin **daešureboda** da lešsa. 'Why would anyone want the woodpecker? No one would hurry to pounce on the skin and carrion [IdO].
26.b. Dro jvirad ğirs, ..., nu davḳargavt..., **davešurot**! 'Time is money..., let's not lose it..., let's make haste.'

27.a. Mosçavleebma adreve icodnen, rom am dğes miğebuli nišani momaval çels simçipis aṭesṭaṭši **čaeçerebodat.** 'The pupils knew in advance that the grade received this day would be entered into their [IdO] high school diploma the following year. (KEGL)
27.b. Raṭom miliciaši ar **čaeçerebi**? 'Why don't you join the police (i.e., sign up for the police)?' (KEGL)

As can be seen from the above examples, there is no information as to the presence or absence of an O_3 in the verb form,[18] and thus we cannot speak about a verbal marking of O_3 here.

Passives of state. Passives of state in the present series can be absolute or occur with the dative object markers *u-* (objective version) or *a-* (superessive). However, these oppositions are totally neutralized in the future and aorist series, where all passives of state take the *e*-prefix. Examples (from Tschenkéli 1958:440-45):

28.

PRESENT SERIES:			FUTURE SERIES	
NEUTRAL	OBJECTIVE	SUPERESSIVE		
sc̣eria	uc̣eria	ac̣eria	ec̣ereba	be written
sdevs	udevs	adevs	edeba	be lying
hqria	uqria	aqria	eqreba	be scattered

It should be clear from the above that whatever the function of the so-called O_3 preradical vowel *e-* in the above examples, it cannot be the signaling of a third person dative object. This is because the marker in question is common to both absolute and relative forms of the verb. Consequently, in the future and aorist series of passives of state there is no verbal marking of an O_3; the only possible indication of an indirect object is the presence of a dative noun phrase in the clause.

Indirect evidence of the presence of O_3. There is a set of circumstances in which the verbal form gives a formal indication of the presence of an O_3, be it direct or indirect. This occurs when 3d person number agreement is with a plural "direct object" or "indirect object." Note such a sentence as:

29. Gamovida da xelis momc̣ert [IdO] **egzavnebat** Didi sabč̣ota enciḳlopediis ṭomebi. 'It came out and the volumes of the *Great Soviet Encyclopedia* were sent to the subscribers.' (KEGL).

Here the form *egzavnebat* unambiguously indicates the presence of a 3d person DAT, but it does this by a marking for **number**, by the suffix *-t*. The *-t* cannot be correlated with S_1 or S_2 (both would have the ending *-i,* not *-a;* in addition, S_1 would have the prefix *v-).* Nor can the *-t* be correlated with O_2 (there is no prefix *g-).* It should be noted, though, that the overt indicator is of number and not of (object) person. With the corresponding singular noun phrase there would be no corresponding unambiguous, formal indication of the presence of a dative object.

Perfect series forms of II. conjugation verbs. The one environment which seems to show an unambiguous formal marker of the presence vs. absence of a third person dative is the perfect series of II. and IV. conjugation verbs. Here, though, it is the very verb form (stem) itself that indicates the presence of a dative (the stem is based on the corresponding perfect participle for absolute forms and on the stem of the verbal noun for relative forms); the presence of a dative noun phrase is not necessarily marked by an objective prefix (and never by a preradical vowel. Oppositions that may be viewed as formally neutralized in other screeves (present/future and aorist for II. conjugation verbs and future and aorist for IV. conjugation verbs) are contrastive here. Examples:

30.	PRESENT/FUTURE	PRESENT PERFECT	
REL.	moekceva	mohkcevia	treat s.o. (well,badly)
ABS.	moekceva	mokceula	reach (a place)
REL.	moešureba	mo(h)šurebia	hurry to
ABS.	moešureba	mošurebula	make haste, rush
REL.	untia, aor. ento	ntebia	sthg. is burning at s.o's
ABS.	antia, aor. ento	ntebula	sthg. is burning

It must be emphasized that what is truly distinctive here (in particular with reference to the marking of O_3) is the very stem of the verb, and not the presence or absence of object markers.

There is something very strange about this environment of contrast. In general, as we know from the theory of markedness, the I. conjugation is less marked than the II. conj., and, of the verbal series (present, future, aorist, perfect), the perfect is the most marked. Consequently, we would expect the (unambiguous) marking of IdO_3 to be most likely in the less marked I. conjugation verbs, including the perfect series (where, in fact, the verb cannot mark an indirect object), and least likely in the most marked perfect series of II. conjugation verbs. But, in fact, we find just the opposite. Unambiguous marking of IdO_3 is strongest in the perfect series of II. conjugation verbs.

Conclusions. There is no formal marking of DO_3 in modern Georgian. Unlike the distinctive absence of formal marker for S_2, where we are probably justified in positing a Ø-prefix,[19] no such arguments exist for DO$_3$. The preradical ("version") vowels *a-* and *e-* similarly cannot be viewed as markers of indirect objects (specifically, of IdO$_3$, i.e., when not preceded by *m-*, *g-*, or *gv-)*. Though there are counterexamples, *u-* generally does serve to mark the presence of an IdO$_3$, and as such is the only really distinctive marker of O$_3$ in modern Georgian.

REFERENCES

Aronson, Howard I. 1970. "Towards a semantic analysis of case and subject in Georgian." *Lingua*, 25:291-301.
_____. 1989. *Georgian: a reading grammar*. Columbus, Ohio: Slavica.
_____. Forthcomimg. "Paradigmatic and syntagmatic subject in Georgian." To appear in a volume of Proceedings from the Fourth International Conference on the Non-Slavic languages of the USSR. Columbus, Ohio: Slavica.
Asatiani, Rusudan. 1986. "Ṗolip̣ersonalizmis jiritadi punkcia kartvelur enebši." *Macne* 1986-2:144-157.
Benveniste, Emile. 1966. "Structure des relations de personne dans le verbe." In: E. Benveniste, *Problèmes de linquistique générale, 1*. Paris: Gallimard, pp. 225-236.
Gamqrelidse, Tamas. 1981. "'Personenhaltigkeit' und 'Valenz' des Verbs." *Georgica* 4:65-70.
GDW = Tschenkéli, Kita. 1965-74. *Georgisch-deutsches Wörterbuch*. Zürich: Amirani.
J̌.m. = Žğenti, Sulịko. 1977. *J̌arisḳacis mama*. Tbilisi: Xelovneba..
J̌orbenaje = J̌orbenaje, B. 1983. Zmnis xmovanṗrepiksuli çarmoeba kartulši. Tbilisi: Tbilisis universiṭeṭi.
KEGL = Čikobava, Arnold, ed. 1950-64. *Kartuli enis ganmarṭebiti leksiḳoni, 1-8*. Tbilisi: Sak. S.S.R. mecnierebata aḳademia.
Lortkipanidse-Piel, Marina. 1986. "Ein Versuch der Funktionsdeutung des georgischen Charaktervokals *e*." *Sprachbau und Sprachtypologie (Wissenschaftliche Beitäge der Friedrich-Schiller-Universität Jena)*, 63-69.
M.ğ. = Dumbaje, Nodar. *Mziani ğame*. Tbilisi: Sabç̌ota Sakartvelo.
Nozaje, L. 1958. "*In-eniani* vnebitis urtiertobis saḳitxis kartulši." *Iberiul-ḳavḳasiuri enatmecniereba* 9-10:129-35.
Šanije, Aḳaḳi. 1973. *Kartulienis gramaṭiḳis sapujvlebi, 1. Morpologia*. Tbilisi: Tbilisis universiṭeṭi.
Sicili = J̌aparije, Bežan. 1971. "Sicili." *Mnatobi* 47-11:103-11.

Suxišvili = Suxišvili, M. 1979. "Uaxloesi obiekṭisa da subiekṭis gaučinareba gardamavali zmnis šemcvel šesiṭqvebaši." *Iberiul-ḳavḳasiuri enatmecniereba* 21:57-67.

Tschenkéli, Kita. 1958. *Einführung in die georgische Sprache, 1.* Zürich: Amirani.

Vendler, Zeno. 1967. "Verbs and times."*Linguistics in philosophy.* Ithaca, N.Y.: Cornall University Press, 97-121.

Vogt, Hans. 1971. *Grammaire de la langue géorgienne.* Oslo: Universitetsforlaget.

NOTES

[1] For the purposes of the present paper, the term *subject* will be initially defined as in Aronson 1970 with the exception that IV conjugation verbs will be regarded as only those verbs without the preradical vowel *e-* in present series screeves and which form their future and aorist series screeves with the preradical vowel *e-*; the direct object is that noun phrase marked by the dative in the present/future series and by the nominative in the aorist and perfect series. The indirect object is marked by the dative in both the present/future and aorist screeves of I, II, and III conjugation verbs.

Terminology used in this paper is basically that found in Aronson 1989. I. conjugation is called elsewhere 'transitive' or 'class 1.'

[2] The reference here is to person alone and not to the combination of person and number. The combination of person and number can be marked prefixally (the 1sg. prefix *m-* and the 1pl. prefix *gv-),* suffixally (e.g., the 3sg. *-s, -a* vs. the 3pl. *-(n)en, -es,* etc.), and by a combination of prefix and suffix (e.g., 2pl. *g-...-t,* 3pl. *u-...-t).*

[3] The question of just what is subject, what is direct object, and what is indirect object in Georgian is one of the thorniest questions in the study of this language. Since a precise definition of these terms is not essential to the present discussion, I shall operate with the definitions of subject, direct object, and indirect object as given for the present, future, and aorist series of I., II., and III. conjugation verbs (see Aronson 1989).

Following these definitions, we get:

Subject marker: prefixes *v-, x-.*
Object marker: prefixes *m-, g-, gv-.*
Indirect object (DAT) marker: prefixes *h-/s-, u-.*

[4] These rules for determination of person(s) marked by a verbal form are further complicated by the presence of suffixes marking number and by their interaction with person/number suffixes. For a complete description, see Tschenkéli 1958:

passim.

[5] The same situation, *mutatis mutàndis,* holds true in the perfect series of I. and III. conjugation verbs, where the desinence *-a, -s* gives no indication that there is in fact a direct object in the clause. Cf. *gaugzavnia* 'he has sent **it**' and *uduğnia* 'it boiled' (without direct object).

[6] See Suxišvil 1986:52-55 for a further discussion.

[7] Abbreviations refer to Tschenkéli's classification of Georgian verb forms found in volume I of the *GDW*:xxxv-xxxviii

[8] I.e., the possibility is limited basically to imperfective forms.

[9] The *a*-series markers also can indicate direct objects; from a form with this series markers alone it is generally impossible to predict on formal grounds whether there is a direct object or an indirect object (or no object at all).

[10] Although there is no 2nd person in the English translation of such forms as *gaxlavar(t)* 'I am (polite),' we should regard the Georgian as clearly having a reference to the addressee.

[11] It is necessary to give an example here with a 1st or 2nd person dative actant because a form marking a third person actant would be homonymous with a form of the type in the (a.) sentence; cf. *daaxvelebs [is]* 'he (NOM) will cough'—*daaxvelebs [mas]* 'he (DAT) will cough.'

[12] For the opposition personal (= 1st and 2nd persons)/non-personal (= 3rd person), see Benveniste1966:228, 231.

[13] The same is true, *mutatis mutandis,* for its occurrence in the perfect series. In the perfect of impersonal verbs such as *c̣vims* 'it is raining,' *tovs* 'it is snowing,' etc., the *u-* does not indicate the presence of a dative actant: *uc̣vimia* 'it apparently has rained,' *utovnia* 'it apparently has snowed.'

[14] Šanije (1973:550-51) regards the *e* here historically as part of the root of the verb.

[15] The view is widely expressed that *e-* serves basically to mark indirect objects. Cf. Lortkipanidse-Piel: "Sobald der CV [Charaktervokal] *e* in der Verbform auftritt, erscheint auch ein indirektes Objekt" (1986:66) and "Der CV *e* jedoch steht unbestritten für das indirekte Objekt" (1986:68) .

[16] For a definition of IV. conjugation verbs which includes passives of state, see Aronson forthcoming.

[17] It is important to note that these verbs are almost always **active** in meaning, usually verbs of motion, and in this respect differ from the majority (though by no means all) of II. conjugation verbs in *i-/e-*, which tend to be passive in meaning. Being active in meaning, they tend to take animate or personal nouns as the grammatical subject. (Although, perhaps, the reasoning is reversed here; it may be better to say that these verbs as a general rule take animate subjects and therefore are active in

meaning.)

[18] But, in verbs of this type, there *is* a formal opposition in the perfect series: absolute *dašurebula*, relative *da(h)šurebia*. See below.

[19] Additionally the Ø-prefix does have, in two verbs, a non-zero variant, *x-*; cf. *x-ar* 'you are' and *ç̣a-x-val* 'you will go.'

PHYLETIC LINKS BETWEEN PROTO-INDO-EUROPEAN AND PROTO-NORTHWEST CAUCASIAN

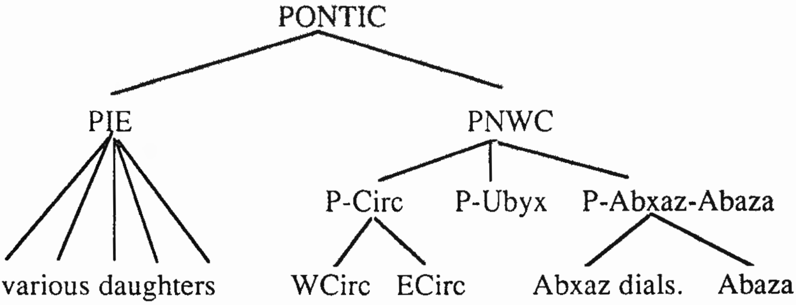
John Colarusso
McMaster University

Introduction. In 1964 Paul Friedrich (1964:209), in a review of Aert Kuipers' work on Kabardian (Kuipers 1960), first made the informed suggestion that Proto-Indo-European (henceforth PIE) might be phylogenetically related to "Proto-Caucasian."[1] Friedrich's suggestion was based on the emerging typological similarities between PIE and some of the Northwest Caucasian languages. The Northwest Caucasian look of PIE, a look which set it widely apart from any of its daughters, had first emerged under the work of internal reconstruction done by Benveniste (1935) and Lehmann (1952). The typological parallels between this early PIE and a Caucasic[2] language were first noticed by Aert Kuipers (1960) for Kabardian and were later taken up by W. S. Allen (1965) when he discussed Abaza vocalism. Kuipers devoted a chapter of his monograph to the parallels between PIE and Kabardian vocalism, which is very similar to the vocalism of Abaza. I myself (Colarusso 1981) have examined typological parallels involving consonantism, particularly matters regarding the so-called laryngeals of PIE and their possible typological correlates among consonants of the Northwest Caucasian languages. Typological parallels betwen PIE and the South Caucasian family, Proto-Kartvelian, were also put forward in the 1960s (Gamkrelidze 1967, 1966; Gamqrelije and Mač'avariani 1965; but note Kuipers 1983), suggesting that at the least PIE formed an areal grouping with the ancient Caucasic languages. In 1987, after I had presented a reconstruction of Proto-Northwest Caucasian (henceforth PNWC) (Colarusso, 1989a), Eric Hamp suggested to me (personal communication) that I endeavor to determine if PIE and PNWC might be genetically related. The following paper presents my first results suggesting that PIE and PNWC are genetically related at a phyletic level.

Proto-Pontic. I shall term the language from which PIE and PNWC may have descended Proto-Pontic, or simply Pontic, after the classical name for the Black Sea, *Pontus Euxinus*, which I assume was near to the homeland. In the past twenty years the archaeological work of Gimbutas (1985, 1980, 1977, 1974, 1973; see also Mallory 1989, ch. 6) has placed the most likely PIE homeland in the Northwest Caucasus. More recently Gamkrelidze and Ivanov (1985, 1984) have argued that it lay just south of the Caucasus, in eastern Anatolia. In either case a phyletic link with a Caucasian family is plausible. My own work in comparative mythology (Colarusso 1989b, 1984) has suggested cultural contacts with the Caucasus at a period of Indo-European unity. Whether or not Proto-Pontic is in fact Proto-Caucasian or Proto-North Caucasian, in other words, whether or not PNWC enjoys a special phyletic link with PIE not enjoyed by other Caucasic language families, rests upon further work in historical Caucasic linguistics. The genetic links between PNWC and Proto-Northeast Caucasian (PNEC) now seem quite plausible (see, for example, Chirikba 1986; Abdokov 1983). Thus, if the present study seems a worthy start, then the reader should be prepared to view PIE as one of an ancient complex of cognate languages centering about the Caucasus. In my opinion, time will show that PIE is closest to PNWC, in fact sharing certain innovations with the northern dialect area of PNWC.[3]

Diagram (1) gives a rough idea of the links that I shall explicitly put forward.

(1) Proto-Pontic



Time Depth and Types of Evidence. For such remote phyletic links as Pontic questions of time depth and types of evidence must be addressed. While it is in principle impossible to establish exact dates based upon linguistic facts alone, I have nevertheless put forward a tentative time frame in (2) which seems to permit room enough for the type of differentiation required for both PIE and PNWC.

(2) Tentative Time Depths

1. 3,000 - 4,000 BC: Comparatively reconstructed PIE
2. 5,000 - 6,000 BC: Internally reconstructed PIE
3. 2,000 - 4,000 BC: PNWC
4. 7,000 - 9,000 BC: Pontic

At such a time depth of nine to eleven thousand years standard cognate evidence will not loom as large as in more conventional reconstructive effort. Accordingly, I shall examine three types of evidence. First, typological parallels (of phonological inventories), suggest not only an areal grouping of PIE and the Caucasus, but also show some strong defects in the PIE inventory, even as revised by Gamkrelidze, Ivanov (1973, 1972, 1967) and Hopper (1982, 1977a, 1977b, 1973). I have made modifications to the PIE inventory which make it far more plausible typologically to provide a basis for correspondence sets. Second, I examine morphological cognates, (compare Goddard 1975). Such morphotactic cognates are strong in the case of nouns, but a bit weaker in that of verbs. One of the strongest sets of data involves the homonymy of morphemes. Indeed the ability of Pontic to explain long-standing homonyms or confusions in morphology within PIE is most striking and is at this stage of work the strongest argument for the cognacy of PIE and PNWC. Odd relict forms within NWC are also explained by Pontic with much more than chance success. Many of these morphological investigations produce transparent explanations of PIE morphology at the level of Proto-Pontic. This is another very powerful argument for the cognacy of PIE and PNWC. Third, lexical cognates can be expected to be few at such a time depth. Nevertheless, a simple search found twenty basic items of good quality, (64)-(83). Many more await the resolution of a few details before they, too, can be published. I turn now to these various categories of evidence.

Typological Evidence. The phonemes of "Classical PIE" are shown in (3).

(3) Classical PIE

p	bh	(b)	m			w
t	dh	d	s	n	r	l
k	gh	g				y
k^w	gh^w	gh				

ə$_1$(E) ə$_2$(A) ə$_3$(O) ə$_4$(A, but not in Hittite in Anlaut);
vowels: e ~ o (plus tonal stress)

Typological arguments based not only upon inherent plausibility, but also upon problems in the development of the PIE system in certain of its branches (Colarusso 1981), have led to a suggested modification in (4) that makes PIE look more like a Caucasic language.

(4) *New-PIE (Gamkrelidze-Ivanov-Hopper), plus palatals*

p^h	(p')	b		m			w
t^h	t'	d	s	n	r	l	
k^{hy}	k'^y	g^y					y
k^h	k'	g					
k^{hw}	k'^w	g^w					

ə$_1$(E) ə$_2$(A) ə$_3$(O) ə$_4$(A, but not in Hittite in Anlaut)
ə ~ a (plus tonal stress)

There are some unrecognized problems with (4), however, that I attempted to point out in an earlier work (1981). First, there are not enough spirants (apart from some of the laryngeals). Second, there are not enough rounded segments for a vertical-vowel system language. Such systems evolve by a rare but natural process in which the features of the syllable core are reassigned to the consonantal syllable periphery. Rounding is one of the most stable of these once so reassigned. Third, we now know enough about the effects and history of the laryngeals that any presentation of PIE needs at least feature specifications for them. Therefore, in (5) I present a typologically more accurate form of PIE, which I term "fortified PIE" and which I shall write in phoneme slashes / ... /.

(5) Fortified PIE (after Colarusso, 1981)

p^h	b	-		m			w
t^h	d	t'	s	n	r	l	
k^{hy}	g^y	k'^y					
(k^h	g	k')					
k^{hw}	g^w	k'^w					
q^h	-	q'	x	ɣ			
q^{hw}	-	q'^w	x^w	ɣ^w			
			ḥ	ʕ			
			ḥ^w	ʕ^w			
		ʔ	h				
		ʔ^w					

ə ~ a (plus tonal stress)

The Laryngeals. In (5) I have given substance to the laryngeals based upon detailed considerations of PIE phonology (Colarusso 1981). I cannot repeat these here, but try to summarize my arguments by an "eightfold way." Any phonologically realistic account of the PIE laryngeals must account for these eight facts.

First, in oldest PIE some true laryngeals produced instances of "inherently" long vowels, schematically shown in (6).

(6) Earliest Laryngeal Loss Giving "Inherently" Long Vowels.
*ē = */əʔ/ (ə_1); *ō = */əʔ^w/, */aʔ^w/ (ə_3); *ā = */əh/, */ah/ (ə_4)
perhaps also:
*ē = */ə-ə/; *ō = */a-a/ [aː] or [ɛː] (contrast: *ā < */əh/, */ah/ = [ɑː])

(parallel: PNWC *h > *a)

Second, at this stage the other segments destined to become "laryngeals" would have persisted as segments without obvious effects into the period of unity. The best candidates for such segments are in (7).

(7) Earliest Persistent "Laryngeals"
/x, ɣ, xʷ, ɣʷ, ḥ, ʕ, ḥʷ, ʕʷ/

Third, with a shift from pharyngeal to true laryngeal in the period of early differentiation the [+Constricted Pharynx] members of (7) would have colored vowels as in (8).

(8) Vowel-Coloring Era
/ḥ/ > /h/, /ʕ/ > /a̯/ (ə̯$_2$), /ḥʷ/ > /hʷ/, /ʕʷ/ > /o̯/ (ə̯$_3$)
(parallels: Abx /ḥ/ > /h/, /ʕ/ > /a̯/, /ʕʷ/ > /yʷ/)

Fourth, once these segments had become true laryngeal glides they were dropped with compensatory lengthening post-vocalically, even throughout Anatolian by means of the natural rule in (9). This would have been a period of early dialect formation.

(9) Rule of Laryngeal Loss in Early Dialect Period.
[-syll, +low] > [+long] / V_____. (phonemes in (8) were lost)
(parallels: Circ /ah/ > [ɑ:], Abx /(ə/a)ʕ/ > /aa̯/ > [ɑ:])

Fifth, in Anatolian some sort of segments persisted in some post-vocalic positions in Anatolian (Hitt *paḫḫur* 'fire,' *meḫur* 'season,' *seḫur* 'urine, filth'). Significantly similar segments gave velars allophones in Italic: Lat. *senātus*, *senex*. These would have been the old velar or uvular spirants, as in (10).

(10) Old Persistent "Laryngeals" of Anatolian
/x, ɣ (ə̯$_{1/2}$), xʷ, ɣʷ (ə̯$_3$)/

(parallels: persist in Circassian and Ubykh)

Sixth, outside Anatolian these same segments were lost with vowel-coloring and compensatory lengthening, as in (11).

(11) Loss of "Persistent Laryngeals" Elsewhere
/x/ > /h/, /ɣ/ > /a̰/, /x^{w}/ > /h^{w}/, /ɣ^{w}/ > /o̰/ (steps 8 and 9 again)
(parallels: similar history in Abx-Abz, e.g. */ɣ/ > /ʕ/ > /a̰/)

Seventh, there is ample evidence that all the laryngeals caused source feature effects. These are of the three types seen in (12). The scheme in Fortified PIE (henceforth simply PIE) neatly accçounts for all of these effects in the simplest way possible.

(12) The Three Source Feature Effects
(a) Glottalization (Voicing): */-p^{h}ʔ^{w}-/ > */-p'^{w}-/ > *-*bo*- : Sk *pībati* 'he drinks,' Ir. *ibim*, Gk πῖνω
(b) Voiceless Aspiration: */-t^{h}ḥ(^{w})-/ > Ind-Iran *-*th*-
(c) Voiced Aspiration: */-t^{h}ʕ(^{w})-/ > Ind-Iran *-*dh*-

Eighth, and last, laryngeals seem to have caused apparently contradictory lowering in some cases but raising in others, as in (13).

(13) Apparently Contradictory Laryngeal Effects
Gk θυγάτηρ, Sk *dúhitā*,
only [+CP] with its low, strong F$_1$ can do both
(parallels: Bzyb Abx /ʕ/ > /a̰/, /ʕ^{w}/ > /y^{w}/, /y/)

This can only be understood if one realizes that this is a pharyngeal "signature" in which an acoustic assimilation produces the opposite effects of an articulatory assimilation (Colarusso, 1985). Pharyngeals have a formant structure with a low and powerful first formant. This gives the impression of a high front vowel. At the same time they are made with tongue root retraction and often with tongue root lowering, which results in approximation of the epiglottis over the adytus (opening of the larynx). Such pharyngeals or "adytals"[4] produce low vowels by articulatory assimilation.

The phonology of the PIE "laryngeals" is complex, but can now be explained by phonological theory and must not be dispelled by elaborate arguments involving leveling and other arbitrary gestures as is now so often the case. One of the few workers who tries to utilize realistic laryngeals and follow them whither they lead is Eric P. Hamp (see, for example, Hamp 1990). This gives many of his reconstructions a distinctly Northwest Caucasian cast.

Morphological Cognates. The form and position of morphological peculiarities can be enormously useful in retrieving ancient phyletic links, so much so that this effect can compensate for the relatively limited phonemic inventory usually associated with morphology. In the present matter there is a good case for nouns (both PIE and PNWC had N-(suffix)^m), and a less strong one for verbs. Pontic seems to have been moderately isolating, much like a NEC language. Subsequent history led the verb to be highly inflected, but in different ways in the two families: NWC: (prefix)^m-V-(suffix)^n, IE: (prefix)^1-V-(suffix)^n. Nevertheless, morpheme cognates are good evidence for two reasons. First, bound morphemes are unlikely to be borrowed outside the forms in which they occur. Second, PIE and PNWC morpheme cognates show a high congruence in otherwise unmotivated homonymy. In some cases PNWC forms can explain peculiarities of PIE inflection.

Sample of Nominal Suffixes. I turn now to an actual presentation of morphological cognates, starting with the noun (and adjective)and treating primarily derivational affixes. Abbreviations in the following are: Bzh = Bzhedukh, (W)Circ = West Circassian, PC = Proto-Circassian, Kab = Kabardian (East Circasian), Ub = Ubykh, Abx = Abkhaz, Abz = Abaza, A-A = Abkhaz-Abaza. I have followed the usual abreviations for the Indo-European languages. Others are V = verb, N = noun, preV = preverbal particle. Each entry is headed by its PIE form, first in its classical representation and then, within parentheses, by its fortified one.

(14) Athematic **-ø, (*/-ø/) :* thematic **-e/o-, (*/-ə, -a/)*
  (a) PIE: Gk *ϝάναξ* ‘lord,’ < */wánakt-s/, vs. λόγο-ς ‘word’ < /log-o-s/
  (b) PNWC: tendency of some languages to produce roots with vowelless allophones or even underlying forms:
  Bzh WCirc /šʰə/ ‘brother,’ /za-šʰ-x̂ʰa-r/ all-brother-pl-abs = ‘the brothers (coll)’; /pq/ ‘bone, frame’ > [pqə], /w-pq-xʰa-r/ > /p-pq-x̂ʰa-r/ your-bone-pl-abs = ‘your body’; Ub /tx̂/ ‘shoulder, back’> [tx̂ə], /á-tx̂/ the-shoulder,back, /tx̂ə-pqə́/ back-bone, /a-tx̂-pqə́/ the-back-bone; vs. Bzh WCirc /psaaƛa-x̂ʰa-r/ word-pl-abs = ‘(the) words’

(15) PIE **-(e)w- (*/-(ə)w/)* in Adjs
  (a) PIE: Gk πολ-ύ-ς ‘much,’ Sk *pur-ùḥ,* Goth *fil-u*
  (b) PNWC: */-u/, */-əw/ ‘predicative’ and ‘adverbial,’ Circ /yən-əw/ big-pred, /psta-w/ all-adv, Ub /ø-dɣa-ø-bya-w-nə/ 3-when-3-see-adv-gerund = ‘when he saw him,’
  (c) Pontic: */-w/.

(16) PIE **-yo- (*/-ya-/)* Abstract Adjs **-iyo- (*/-iya-/)* (see ‘collectives,’ (30))
  (a) PIE: Sk *gáv-ya-ḥ* ‘bovine,’ *áśv-iya-ḥ* ‘of the horse, horse-like,’ *árya-* ‘Aryan’
  (b) PNWC: */-ĝa/ > Circ /adə́ĝa/ ‘Circassian,’ Abz /-rʕa/ ‘people,’*/-ya/ > WCirc /da-a-ya/ nut-con-one of = ‘nut tree’ [-iye] vocalization of /-ya/ common in Ub and A-A,
  (c) Pontic: */(rə)ĝa/ ‘people’ (see (30)),
  */-ya/ ‘the one of,’ adjectival suffix.

(17) PIE **-yo- (*/-ya-/)* opposition with other terms
  (a) PIE: Lat *alius* ‘the other,’ Gk δεξ-ιό-ς ‘the right one,’ Goth *niu-ji-s* ‘the new one,’
  (b) PNWC */-gʸa/ ‘and’: Ub /-g’a/ ‘and,’ Circ /-əy/ ‘and’ (of clauses), Abz /-gʸ-/ ‘and’ (preV),
  (c) Pontic */-ge/ ‘and’ (of pairs).

(18) PIE **-en- (*/-ən-/)* used in oblique cases
  (a) PIE: Goth *guma* 'man,' *gumin-s* 'gen.,' Lat. *homō*, *homin-is* id.,
  (b) PNWC: */-n/ or */-m/ oblique case, genitive formation: Circ /ƛ'ə-m ə-qʷ/ man-obl his-son,
  (c) Pontic: */-m/, (rather than */-n/, because the former is typologically more marked so the shift */m/ > */n/ may be explained as a typological simplification).

(19) PIE **-no- (*/-na-/)* secondary NPs
  (a) PIE: Lat *Lūna*, Praenestinian *losna* < **lowks-no-*, Av *raočaḥ* 'light, lamp'; Sk *pūr-ṇá-ḥ* 'something full,'
  (b) PNWC: */-nə-/: frozen derivational suffix in Circ: Bzh /č'aš'-nə-nəqʷa/ night-/nə/-half = 'midnight,' so-called "syllabified connective" in /λa-nə́-sʰtʰa/ 'scissors,' /ś'a-nə́-ɣa/ know-/nə/-ness = 'knowledge' (so, by this last form/-nə-/ cannot be an old genitive),
  (c) Pontic: */-na-/, */-nə-/.

(20) PIE **-eno- (*/-əna-/), *-ono- (*/-ana-/)* participle in Germanic
  (a) Gmnc: Goth *itan* 'eaten,' *bit-an-s* 'killed,'
  (b) PNWC: Abz /-ən/ "pro-tense", replaces tense in concatenated or subordinated ("dependent") forms: /s-č'a-n/ I-eat-dep; Ub /-nə/, /-na/ old gerund, /a-la-sə-nə ... ø-dɣa-ø-bya-w-na ... ø-yə-q'a-q'a/ she-there-sit-ger ... him-when-she-see-adv-ger ... it-she-say-past = 'she was sitting there ... when she saw him ... [and] she said,'
  (c) Pontic: */-əna/ old participle ending.

(21) PIE **-(t)er (*/-(tʰ)-ər/)* old kinship suffix
  (a) PIE: **swesor* > Lat. *soror* 'sister,' E *sister*, Arm *kʰoyr*, Pers *xʷāhar*, **p(ə)tə́r(s)* > Gk πατήρ, Sk *pitār*, Lat. *pater*, Arm *hayr*, Ir *athir*, Goth *fadar* 'father,'

(b) PNWC: */X-tʰ-ər/ X-be-part(iciple) = ‘the one who is X,’ *X-ər X-pt = ‘the one who is X,’ ‘the X’

(c) Pontic: */-tʰ-ər/ -be-part, */-ər/ -part.

(22) PIE **-er (*/-ər/)* in nom-acc, sg, neut., **-en* (*/-ən/) in obliques

(a) PIE: Sk *ū́dhar* ‘breast,’ *ú–dhnaḥ* gen.

(b) PNWC: */-ə/r in abs(olutive) (if neuter [-agentive], one will not have an ergative role), */-əm/ or */-ən/ in obl(ique) cases: Circ /ƛ’ə-r/ man-abs, /ƛ’ə-m/ man-obl; Ub /tə́t/ ‘man(abs),’ /tə́t-ən/ man-obl

(c) Pontic: */-ər/ abs., */-əm/ obl., (note (17, c)).

(23) PIE **-yes-/*-yos-/ (*/-yəs-/* or **/-yas-/)* Comparative*; *-i-s-t(h)o-* (*/-y-s-tʰa-/ or */-y-s-da-/*)* Superlative

(i) Comparative:

(a) PIE: Sk *svá–d-īyas-* ‘sweeter,’ Gk ἡδΐω id.

(b) PNWC: */-y-ćʰ/ -dir(ection)-be excessive > Bzh WCirc /-šʰʸ/ ‘excess,’ Ub /ċa-/ comp;

(ii) Superlative:

(a) PIE: Sk svá–d-iṣṭha-ḥ, but Gk ἥδ-ιστο-ς (*ἥδ-ισθο-ς odd!)

(b) PNWC: superl = comp + ‘exactly’ */-y-ćʰ-(də)da/, whence > */-y-ćʰ-tʰa/ > PIE *-y-s-to, or */-y-ź-da/ > PIE *[-i-z-dho] (*/-i-s-dho/)

(c) Pontic: */(-y-a-)ćʰ/ (-dir-dat(ative)-) be excessive (PIE *-yas), */(-y-)ćʰ/ (-dir-) be excessive (whence the Circassian form), */ćʰ-a/ be excessive-dat (whence the Ubykh form).

(24) PIE **-ter- (*/-t'ʰər-/), *-tel- (*/-t'ʰəl-/)* Agents

(a) PIE: Gk γενε-τήρ, γενε-τώρ, OCS *bljustelji̮* ‘observer’; NB: Hitt has only **-l-*.

(b) PNWC: Abz /-la-/ instrumental, /q̄ac’a-la/ man-instr = ‘by means of the man,’ in the north this is /-r(a)-/: Kab /wa-r-k’ʸa/ you-instr-instr = ‘with you(r help)

(c) Pontic: */(-t^{h}ə-)l-/ instrumental. PIE */-t^{h}ə-/ is probably an innovation based upon the extension of the genitive as an oblique case (cf. Abz /q̄ac'a-ta/ man-gen = 'of, from the man'); note, part of PIE also shares an isogloss (*/l/ > */r/) with northern PNWC.

(25) PIE **-tro- (*/-t^{h}ra-/), *-tlo- (*/-t^{h}la-/), *-dhro- (*/-dra-/), *-dhlo- */-dla-/)* Instrumentals

(a) PIE: Sk *mán-tra-ḥ* 'prayer,' Lith *(pa-)men-klas* < **men-tla-* 'monument,' Lat *pō-culum* < **pō-tlo-m* 'drinking cup,' OIr *cé-tal* < **kan-tlo-* 'song,' OHG *sta-dal* < **sta-þla* 'barn,' Gk γένε–θλο–ν, ἄρο–τρο–ν 'plough,' Lith *ár-kla-s* id., Czech *rá-dlo* id.

(b) PNWC: */-la-/ (same as (23)), Abz /ḥaq̄^{w}-la/ rock-instr = 'with he rock'; Circ /zə-rə-z/ one-by-one, Bzh WCirc /ø-z-a-r-a-ƛaɣ^{w}ə-ɣa-x̂^{h}/ 3-reciprocal-dat-instr(umental)-dat-see-past-pl = 'they saw one another'

(c) Pontic: */-t^{h}a-la-a/ -gen-instr-dat (like Circ reciprocal) > PIE*/-t^{h}la-/, */-dla-/ (with assimilation), or */-t^{h}ra-/, */-dra-/ in more northerly form.

(26) PIE **-men- (*/-mən-/)* nominal action affix

(a) PIE: Sk *bhár-ma, bhár-ī-man-* 'action of carrying,' Gk φέρ-μα,

(b) PNWC: Kab /wə-ma/ strike-/ma/ (old affix) = wooden club for hammering

(c) Pontic: */-m(a)n-/ .

Other Endings. I turn now to some other endings, such as participles, abstracts, cases and such.

(27) PIE **-ent, *-ont, *-n̥ (*/-ənt/, */-ant/, */-nt/)* Active Participle

(a) PIE: Lat. *dēns*, *dentis* (gen.) 'tooth' (lit. 'the eater'); Gk ὀδούς, ὀδόντος (gen.)); Lith *dantis*, Goth *tunþus*,

(b) PNWC: Abz /-n/, Ub /-nə/, /-na/ old participles, plus Circ /-t^h/ durative (distributed) tense,

(c) Pontic: */-(a)n-tʰ-/ ‘participle-durative.’

(28) PIE _*-welos (*/-wəlas/), *-welot-(*/-wəlatʰ-/)_ Perfect Active Participle

(a) PIE: Gk -(Ϝ)ός neut. nom-acc, (Ϝ)οτ-ος gen.

(b) PNWC: */-w(a)-/ aspect sfx, Kab /-w-/ progressive asp, /s-a-w-šx̂/ I-pres-prog-eat = ‘I am eating,’ Abz /-w(a)-/ id., /s-č’(a)-w(a)-n/ I-eat-prog-past = ‘I was eating,’ Abz /-w(a)-z-t-ən/ of dependent past durative, /s-č’(a)-w(a)-z-t-ən/ I-eat-prog-past-dur-dep = ‘that I was eating (for a period of time)’

(c) Pontic: */-wa-z-tʰ-/ > PIE */-wastʰ-/ > */-wos-/, */-wot-/, by dialect splitting.

(29) PIE _*-ā, *-y-ā (*/-əə$_2$/, */-y-əə$_2$/)_ Feminines and Abstracts

(a) PIE: a long scholarly history examining the homonymy of feminines and abstracts,

(b) PNWC: */-xa/ ‘woman’ > Ub /xa-ɣʷa/ you-sfx = ‘you (free woman)’; */w-xa-śəmc’á/ > Bzyb Abx /(á-)ḥʷ(ṡsa)/, Ashxarwa Abx /ḥʷ(sə́sa)/
PNWC: */ʔa/ ‘hand’ > PC /q’a/ (N, V) > */-qa-/ (preV), */-ɣa/ (N-sfx) ‘hand’ or ‘belonging to,’ ‘being in hand,’ or ‘-ness’ (= abstract suffix)

(c) Pontic */-xa/ ‘feminine’ and */-q’a/ abstract suffix have coincided in PIE.

(30) PIE _*-yā (*/-yəə$_2$/)_ Collectives

(a) PIE: Gk φρᾱτριᾱ́, OCS _bratrĭja_ ‘fraternal groups’

(b) PNWC: old collective in Abz /wa-ʕa/ man-coll < */ʕʷa-ʕa/ (cf., Abz /(a)ʕʷə/ ‘man,’ Bzh WCirc /(šʰə-)wə/ ‘(horse-)man,’ Ub /wədə/ devil (< */wə-də/ man-derivational sfx), < PNWC *ĝu-, *w-ĝə- ‘man.’ Whence also Abz /-ʕa/ ‘coll’ (< */-ĝa/) and the /ĝa/ in PNWC */rəĝa/ ‘people,’ Circ /adəĝa/, Ub /a-dəĝá/ ‘Circassian,’ Abz /-rʕa/ ‘people,’

(c) Pontic */-ĝa/ > PNWC */-ĝa/ ‘man(kind),’ ‘collective,’ Pontic */-ĝa/, /-aĝ/ > PIE */-yaɣ/(*-yəə̯$_2$) > /-yā/, by levelling.

(31) PIE Cases

	PIE	PNWC
acc	*-m/-n	/-m/ (obl in Circ)//-n/ (obl in Ub)
gen/abl	*-(ə/a)s (athematic)	*/-š /(old genitive)
gen	*-o-s(y)o (thematic)	*/-š-y-a/ > /-šʸ/ obl of pronouns in WCirc
abl	*-ō (thematic)	Ub /-xʸa/, A-A /-xʸa/ ‘place’ or */a-a/ vowel-in, as with final /-a/in Circ /-y-a-pƛ-a/ -3-dat-look-in
dat *	-əy-	/-y-(a-)/ dir-(dat-), Circ preV
loc	*-i	Circ preV /-y-/ direction, old Bzh
WCirc		dat of pronouns /-y/
inst	*-ē, *-ō	*/-ə-a/ > */-ə̄/ (?), */-a-a/ > */-ā/, with *-a the same as in the thematic ablative

Pontic		
acc	*/-m/	‘oblique case’
gen	*/-š(-y-a)/ or */-y-š-a/	‘old oblique of pronouns or old genitive’
abl	*/-y-(a-)x̂/	-dir-(dat-)place
dat	*/-y-a/	-dir-dat
loc	*/-y/	-dir
inst	*/-a/	-dat

(32) Demonstratives
- (i) *anaphora*
  - (a) PIE: */s-a/ nom, sg, */tʰ-a/ oblique,
  - (b) PNWC:*/śa/ ‘what,’ */tʰə/ ‘where,’ Bzh WCirc /śa/, /śəd/ ‘what,’ /tʰəda/ ‘where,’
  - (c) Pontic: */ś-a/ what-dat , */tʰ-a/ where-dat.
- (ii) deixis
  - (a) PIE : */-w-/ > Sk *asau*,

(b) PNWC: */wə-/ 'that (near hearer),'
(c) Pontic: */wə-/ deixis (near hearer).
(iii) relative
(a) PIE: */ya-/,
(b) PNWC: */yə-/, Bzh WCirc /yə-/ optional absolutive verbal index, Abx-Abz /y-/ relative initial verbal index,
(c) Pontic: */y-a-/ old relative particle-dat.

(33) Personal Pronouns

		PIE		PNWC
		nom	*obl*	
sg	1	*egō̆ (*/ʔək'-ʕʷ-/)	*(e)m-	*/m-/ 'that near me' < Pontic */ʔə-k'-/, */ʔə-m-/,
	2	*tu (*/tʰw/)	*tew/*tw-/*t-	*/w-/ < Pontic */tʰw-/, cf. A-A */b-/ 'you (fem)' < */tb-/ < Pontic */tʰw-/ by regular A-A sound developments
		nom	*obl*	
pl	1	*ways	*nās/*nas	(recent innovations in NWC: Bzh WCirc /t-/, Ub /šʸ-/, A-A /ḥ-/.)
	2	*yus	*wās/*was	PNWC *su-, *w-sə, WCirc /śʷ-/,
		Hitt šumeš, OIr swés		Ub /sʷ-/, Bzyb Abx /ṡʷ-/

Pontic */swə/ > PIE **swə*, is shaped by 2nd sg, but **swə* > late PIE **wōs*/**wos* is shaped by 1st pl .

Preverbs (old nouns). Remarkably the preverbs show some strong parallels between PIE and PNWC.

(34) PIE **perə̯- (*/pʰər-ʔ-/)* ‘before’
(a) PIE: Loc: *per-ə̯-i > Gk πέρι
Gen-Abl: *pr-ə̯-o- > Gk πάρος
Inst: *prō-, *pro- > Lat *prō-*, *pro-*
(b) PNWC: */pʰa-r-(a-y-)/ front-along-(dat-dir-); E *‘ford’* is usually grouped here as a verbal form, but cf. Bzh WCirc /-px̂ə-rə-pšə-/ -through-along-crawl- = ‘to crawl through something (such as underbrush),’
(c) Pontic */pʰəx̂ə-rə-/ through-distributed > PIE **perə̯* (with metathesis of **-x̂-r-);* Pontic */pʰə-rə-/ front-distr > PIE **per(ə̯)-*.

(35) PIE **en- (*/ʔən-/)* ‘interior’
(a) PIE: Loc: *en-i > Gk ἔνι, ἔν, Goth *in-*
(b) PNWC Abz /-n-/ in /n-c’a-ra/ in-place-inf = ‘to place inside’; PNWC: */ʔ(a)-/ > Ub /q’á/ ‘hand’; WCirc /-q(a)-/ preV denoting “action in hand”; A-A /-q’a-c’a-/ -hand-set- = ‘to do.’
(c) Pontic: */(ʔə-)n-/ (hand-)in-.

(36) PIE **et- (*/ʔətʰ-/)* ‘without, outside’
(a) PIE: Loc: Gk ἔτι
with deictic */w-/, Goth *ut-*, Sk *ut-*
(b) PNWC Abz /-t-/ ‘from inside out,’ ‘from below upwards,’ /t-ga-ra/ out-drag-inf = ‘to drag something out,’
(c) Pontic: */(ʔə-)tʰ-/ (hand-)out-.

(37) PIE “final **s*”
(a) PIE: Dor Gk ἐνς (Att εἰς), Goth *ut-*, *us-*,
(b) PNWC: old oblique in */-š/,
(c) Pontic: */-š/ old oblique on nominal ancestors of preverbs.

Particles. Particles are so short as to make comparative study extremely difficult, but even here two forms show such close parallels between PIE and PNWC that they can be taken back to Proto-Pontic.

(38) PIE **r* 'and'
  (a) PIE: Gk ἀρ, ρ῾α, ᾰ́ρα, Lith *ir*
  (b) PNWC: */-ra/, Circ /-ra/ 'and'
  (c) Pontic: */-ra/.

(39) PIE **ge (*/k'ə/)* 'because,' 'terminus'
  (a) PIE: Gk γϵ, Hitt *-k*, Goth *mi-k* 'to me,' *au-k* 'because' ('from that'),
  (b) PNWC: */-y-k'/ -dir-instr, PC */-k'ʸa/ > WC /-k'ʸa/, /-gʸa/, /-č'ʸa/,
  (c) Pontic: */k'ə/ 'because, arising from, issuing from.'

Verbal Desinences (change vowel grade of stem) and Suffixes. Even though the subsequent history of the verb in PNWC tended toward massive prefixation and that of PIE tended toward suffixation, there are numerous parallels between the two families so that a strong case for a Pontic verb can be made.

(40) Athematic:Thematic
  (a) PIE: athem. Sk *ád-mi* 'I am eating' : them. *rod-ā-mi* 'I am crying'
  (b) PNWC:
    (i) basic verbs athem (?) */-tʰ-/ 'to be,' */-w-k'-/ -valence-kill-, Ub /ø-s-k'ʷ-q'á/ it-I-kill-past = 'I killed it'
    (ii) verbs with stem-final /a-/ showing thematic conjugation: WCirc /psaaƛa/ 'word,' /t-zara-psaƛa-a-ɣa/ we-reciprocal-converse-th v-past = 'we talked,'
  (c) Pontic: *CVC-afx* forms vs. *CVCa-a-afx* forms with thematic vowel.

(41) Intensive Reduplication
  (a) PIE: Sk *dediṣ-te* 'he teaches and teaches,' OCS *gla-gol-ją* 'I speak'

(b) PNWC: WCirc /-śa-śa-/ -fall-fall- = 'to fall (as of leaves)' (old, "athematic"), /-λa-λa-/ -hang-hang- = 'to dangle,'
(c) Pontic: *CVC- > CV-CVC-*.

(42) PIE themes with *-*ē*-, *-*ō*-, *-*ā*-
(i) (a) PIE: **men-* (*/mən-/) 'to have in spirit,'
(b) P-A-A: */-ma-/ 'to have, to do' (now only in prohibitive form),(c) Pontic: */-mən-/, */-man-/,
(ii) (a) PIE: *-*mn*-ē- (*/-mn-əʔ-/) stative sense: OCS *mĭněti* 'he thinks,' Gk μανῆ-ναι 'to be maddened,'
(b) PNWC: */-q'a-V-/ -horizon-V- = 'V that is of interest to the speaker,'
(c) Pontic: */-ʔa-V-/, */-V-əʔa-/ 'in hand' affix for action of intimate concern to the speaker,
(iii) (a) PIE: *-*mn-ā*- (*/-mn-əə̯$_2$-/) iterative = 'to recall,'
(b) PNWC: */-x-/ 'iterative,' Abz /n-c'a-x-ra/ in-place-again-inf,
(c) Pontic: */-mn-əx-/,
(iv) (a) PIE: *-*mn*-ō- (*/-mn-əə̯$_3$-/): Gk ϝαλῶ-ναι 'to be taken,'
(b) PNWC: ?*/-q'ʷa-/ 'excess,' WCirc /-šx̂ə-ʔʷa-/ eat-too much.

(43) PIE *-*eyo*- *(*/-əya-/)*, *-ī- *(*/-yə̯-/)*, *-*y*- *(*/-y-/)* Causative / Iterative
(a) PIE: Ved *sād-áya-ti* 'he made him sit, he sat him down' ("inherently" long vowel pattern),
(b) PNWC: Ub /-aay-/ 'again,' 'finally' (NB: /aa/ [a:] perhaps involved with root lengthening in PIE),
(c) Pontic: */-aya-/, */-əya-/ 'iterative,' 'resultative.'

(44) PIE Sigmatic Aorist, *-*s*-
(a) PIE: Ved *vém̩-s-i* 'I have won,' Gk ἔπαυ-σ-α 'he has stopped,'
(b) PNWC: */-z-/, Circ /-z-/ stative or accomplished past particle with past pt, Bzh Circ /fa-d-əy-z/ for-be like-past, pt-completely = 'he was completely like him'; Abz /s-č'(a)-

w(a)-z-t-ən/ I-eat-prog-past-distr-dep = ‘that I was eating (for an interval),’ and other forms,
(c) Pontic: */-z-/ ‘past ending of full effect.’

(45) PIE *n-Infix Presents *(CVC-C- > CC-nə-C-)*
(a) PIE: Hitt *ḫar-k-* ‘perish, be destroyed,’ *ḫar-ni-k-zi* ‘he destroys,’ *ḫar-ni-k-anzi* ‘they destroy,’
(b) PNWC: Ub /-n/ dynamic present /ø-fa-ø-bżatʷə-n/ it-down-he-hang-pres = ‘he is hanging it,’
(c) Pontic: */-n-/ n-infix dynamic present.

(46) PIE Primary Active 3rd Plurals in *-*n*-
(a) PIE: 3rd, sg *-*ti* (*/-tʰi/), 3rd, pl *-*(e/o)-n-ti* (*/-(ə/a)-n-tʰi/),
(b) PNWC: Ub 3rd, pl /-na-/, /ø-fa-ø-bżatʷə-na-n/, them-down-he-hang-pl-pres = ‘he is hanging them,’
(c) South Caucasian: Old Georgian /km-n-na/ make-pl-3rd, past = ‘he made them,’[5]
(d) Pontic: */-na-/ third person plural infix of actives.

(47) PIE Middle Voice in *-*dh*- *(*/-d-/)*
(a) PIE: Dor and Hom Gk ἔσ-θ-ω < *ἔδ-θ-ω ‘I am eating,’ (Sk *ád-mi*), Goth *wal-d-a* ‘I dominate,’ OCS *vla-d-ǫ* ,
(b) PNWC: Abz optative of self-interest /s-č’a-n-da/ I-eat-dep-middle = ‘O, if I could eat!,’
(c) Pontic: */-da-/ self-interest forms.

(48) PIE Perfects in *-*k*- *(*/-kʰ-/)*, *-*g*- *(*/-k’-/)*, *-*gh*- *(*/-g-/)*
(a) PIE: Gk τμή-γ-ω ‘I cut,’ τετμή-κ-α perf., νή-χ-ω ‘I swim,’ Att Gk ἔϑη-κ-α ‘he placed it,’ Phrygian αδ-δα-κ-ετ ‘he has made it,’
(b) PNWC: */-q’a/ past, Ub /-q’a/, WCirc /-ɣa/, ECirc (Kab) /-aɣ/ > [A:],
(c) Pontic: */-q’a/, */-ɣa-/ with dialect variation just as in NWC today.

(49) PIE Optative in **-yē- (*/-yəʔ-/), *-yə- (*/-yʔ-/)*

(a) PIE: **es-* (*/ʔəs-/) ‘to be,’ Sk *ás-ti* ‘he is,’ */ʔs-yəʔ-tʰ/ > Sk *s-yā-t* ‘may he be,’

(b) PNWC: */-əy/ optative, concessive, Kab /ø-x̂ʷaaba-ma-əy/ 3-warm-if-even = ‘even if it be warm,’

(c) Pontic: */-yəʔ/ ‘optative,’ ‘even.’

(50) Primary, Active, Present, Athematic **-i (*/-y/)*

(a) PIE:	1st, sg */-m-i/	1st, pl */-məs-i/
	2nd, sg */-s-i/	—
	3rd, sg */-ntʰ-i/	3rd, pl */-(ə/a)ntʰ-i/

(b) PNWC: */-y-/ present, Abz dynamic /s-ʕʷə-y-t’/ I-write-pres-def = ‘I am writing,’ /s-ʕʷə-t’/ I-write-def = ‘I wrote,’

(c) Pontic: */-y-/ ‘active present affix.’

(51) PIE Relic Impersonals in **r*

(a) PIE: 3rd, pl: Sk *śe-re*, Av *sōi-re* ‘they are lying down,’ Brythonic impersonal: Armorican Breton *new gueler* ‘one does not see me,’ Passive: OIr *berir* ‘he is carried,’ Umb *ier* ‘one goes,’ Lat *ī-t-ur* ‘one goes,’ Middle: Tokh B *kal-t-r* ‘he stops,’

(b) PNWC: */-ra/ optional present, Kab 3rd, pl (occasional impersonal nuance) , /ma-a-k’ʷə+a(-r)/ 3-pres-go+intrans(-pres) = ‘they are going’; interrogative force in non-affirmatives /ø-y-a-gʸa-ra/ he-it-dat-read-pres = ‘is he reading it?,’ /ø-y-a-gʸa-r-q’əm/ he-it-dat-read-pres-not = ‘he is not reading it’ (cf., /ø-y-a-gʸa-ś/ he-it-dat-read-affirmative = ‘he is reading it’); Shapsegh WCirc 3rd, past, intrans /rə-k’ʷə+a-aɣ/ 3-go+intrans-past = ‘he went,’ A-A 3rd, pl non-initial verbal index /-r-/, /yə́-q’a-r-c’a-t’/ it-hand-they-set-def = ‘they did it.’

(c) Pontic: */rə-/ ‘third, plural, indefinite person,’ */-ra-/ ‘non-assertive present.’

(52) PIE "*s*-Movable"

(a) PIE: */sp^{h}-/˜ */p^{h}-/ : Sk *spáś-* 'to spy,' *páśyatī* 'he sees'
*/st^{h}-/˜ */t^{h}-/ : Goth *stauta* 'I strike,' Sk *tudáti* 'he strikes'
*/sk^{h}-/˜ */k^{h}-/ : OHG *skeran* 'to shear, clip,' Gk *κεῖρω* 'I shear'
*/sm-/ ˜ */m-/ : OHG *smelzan* 'to melt,' Gk *με̃λδω* 'I melt,' OHG *malz* 'malt'
*/s-/˜ */w-/ : Gk ἕλκω < *σέλκω 'I drag, pull,' Lat *sulcus* 'furrow' (< **solkos*), Lith. *velkú*, OCS *vlěkǫ* 'I pull' < PIE */swəl-k^{h}-/

(b) PNWC: */(-y)-ƛ^{h}-/ > PC */-y-ś^{h}-/ -dir/3-deixis- > PC */-š^{hy}-/ 'there,' entirely optional on verbs, Ub /-la-t^{w}-/ '-deixis-be-' = 'to be there, exist,'

(c) Pontic: */-ƛ^{h}-/ 'there,' (deixis on verbs).

(53) Personal Endings: not much, but note:

(a) PIE: *s-loss: Gk ἡδῑ́-ω 'sweeter' < **swed-íyo-s*; Av *mą-jrō* 'prayer' < **man-tras*, Gk πατήρ 'father' < **pət-ér-s*,
PIE: thematic, 1st, sg, primary, active, present *-ō (*/-ā/) < *-*o-s* (*/-a-s/?),

(b) PNWC: */-s-a-/ -I-pres (active)-, Bzh WCirc /s-a-tx̂ə+a/ I-pres-write+ intrans = 'I am writing,'

(c) Pontic: */-a-s/ 'thematic vowel-first person.'

(54) Futures in *-*(ə̥)s(y)e-/*-(ə̥)s(y)o- (*/-(ə̥)s(y)ə-/* or **/-(ə̥)s(y)a-/)*

(a) PIE: Sk *vak-ṣ-yā́-mi* 'I will speak,' Gk λείψω 'I will leave,'

(b) PNWC: */-š-/ > Abz /-š-/ fut, /s-č'(a)-w(a)-š-t'/ I-eat-fut-def = 'I will eat,' */-x-š-/ > Abz stative futures, /s-bzəy-x-w-š-t'/ I-good-afx-prog-fut-def = 'I shall be good,'

(c) Pontic: */-š-/ '-future-,' */-x-š-/ '-stative-fut-.'

(55) Intensives in *-*sk(e/o)-* (**/-sk^{h}(ə/a)-/)*

(a) PIE: Hitt. endings *-ški-z-i* '-intensive-3, sg-present,' *-šk-an-z-i* '-intensive-pl-3-pres,'

(b) PNWC: */-šx̂o/ > PC */-šx̂ʷə/ > Shapsegh WCirc /-fʷə/, Natukhay Circ /-ŝʷx̂ʷə/, Bzh WCirc /-škʷə/, Kab /-šx̂ʷə/, Confined to nouns, but note other adjectives, such as /ba/ 'much,' that can play adverbial roles: Kab /sə-q'a-mə-k'ʷ+a-žə-fa-nə-w-ta-ba/ I-hor-not-go+intransitive-back-able-fut-def-irrealis-much 'I shall not be able to go back again then, even so!'

(c) Pontic: */-śx̂ə/ > PIE */-skʰə-/ (with special cluster development, as seen also in Circassian).

(56) The Augment **e- (*/ ʔə-/)*

(a) PIE: */ʔə-/ marks the past, as in Ved Sk *á-bharat* 'he carried,' Hom Gk ἔ-φϵρϵ, but it attracts stress as though it were orignally a word, as in Gk παρ-έ-σχον (*πάρ-ϵ-σχον),

(b) PNWC: */ʔ(a)/ > PC */q'(a)/ > Bzh WCirc with preV loss of ejective feature /ø-qə-w-a-s-tʰə-γ/ it-hor(izon of interest)-you-dat-I-give-past = 'I gave it to you' (accomplished transfer of ownership expressed through /-qə-/), Abx /-q'a-c'a-/ '-hand-set-' = 'to do.'

(c) Pontic: */ʔ(a)/ '(in) hand,' originally an independent adverb before the verb denoting accomplishment of action. The development in PIE suggests links between it and northern (Proto-Circassian) PNWC.

Stem Formation *(à la* **Benveniste***).* One of the oldest patterns in PIE is that of vowel-loss in roots or stems as suffixation proceded: C_1VC_2-C_3-: C_1C_2-VC_3-: C_1C_2-C_3-VC_4 (Benveniste 1935). Parallel to this is the vowel reduction pattern of Circassian morphemes in pre-root position in verbs, as in (57).

(57) Pre-Root Vowel-Reduction in Bzhedukh West Circassian

(a) /wə-*qa*-s-ƛaγʷə-γ/
you-hor-I-see-past
'I saw you.'

(b) /wə-*q*ə-ø-ah-*da*-s-ƛaɣ^{w}ə-ɣ/
you-hor-3-pl-with-I-see-past
'I saw you with them.'
(c) /wə-qə-ø-ah-*d*ə-s-əy-ɣa-ƛaɣ^{w}ə-ɣ/
you-hor-3-pl-with-I-he-cause-see-past
'He showed me you together with them.'

If the pattern in (57) is old and is any way related to the PIE patterns, then in some verbs one might expect C_1VC_2- to be prevebal components, while C_3 proved to be a root. In the conventional view one should expect etymologies for C_3 as suffixes to a root. Etymologies for C_3 have proven to be hard to find (though not for C_4). Taking the PIE and Circassian patterns to be related one might look for cases, therefore, in which C_3 proved to be the root. In (58) and (59) there may be just such a pair (Benveniste 1935:151).

(58) **tér-ə$_1$*- (*/t^{h}ə́r-ʔ-/) : Gk τέρ-ε-τρον 'borer' vs. **tr-éə$_1$*- (*/t^{h}r-ə́ʔ-/) Gk τρή-σω 'I bore'

(59) **tér-ə$_2$*- (*/t^{h}ə́r-ḥ-/) : Hitt *tarḫ*- 'to conquer' vs. **tr-éə$_2$*- (*/t^{h}r-ə́ḥ-/) Lat *trāre* 'to cross upon,' *-trāns* 'across'

It is hard to imagine what root */t^{h}ər-/ in conjunction with what enlargements would produce the resulting meanings in (58) and (59). If the first morpheme is not a root but rather a preverb, however, while the enlargements are in fact distinct roots, then (58) and (59) would not only present a plausible situation, but would find straightforward cognates in PNWC, (60)-(63).

(60) PNWC */-t^{h}ə-rə-w-ʔə-/ -surface-distr-valence-stick- = 'to stick into a surface,' WCirc /-t^{h}(-y-a)-ʔ^{w}ə-/ -surface(-dir-dat)-stick- = id.

(61) Pontic */-t^{h}ə-rə-ʔə-/ -surface-distr-stick- > PIE */t^{h}ə́r-ʔ-/, */t^{h}r-ə́ʔ-/.

(62) PNWC */-tʰə-rə-ḥa-/ -surface-distr-enter- = ‘to enter on something or someone,’ ‘to conquer’ (NB: PNWC has the same range of senses for this form as PIE), WCirc /-tʰ(-y-a)-ḥa-/ -surface-(dir-dat)-enter = id.

(63) Pontic */-tʰə-rə-ḥa-/ -surface-distr-enter- > PIE */tʰə́r-ḥ-/, */tr-ə́ḥ-/.

Many of the odd “homophonous” roots or semantically skewed derivations of the sort of (60) and (61) may be amenable to a solution of this type. Further work in this area promises to reveal some of the more obscure cognates between these two families as well as to throw light upon some of the more difficult laryngeal developments within Indo-European history.

Conventional Cognates. In the following I conclude this study with a list of some of the best and simplest cognates of a conventional sort. While they do not bulk large in this study because of the time depth for Proto-Pontic, they nevertheless can be found. Many are of a very striking and forceful character, both phonologically and semantically. In these I give first the Pontic reconstruction, followed by the PIE and then the PNWC histories.

(64) ‘fire’ (‘that which descends (from heaven),’ i.e. ‘lightning’)
  (a) Pontic: */pʰa-x̂ʷə-r/ down-fall-abs/ger = ‘that which falls,’ */pʰa-x̂ʷə-n-i/ down-fall-obl-dat = ‘in the fire,’
  (b) PIE: */pʰaxʷə-r/, Hitt *paḫḫur* ‘fire’ (nom-acc), *paḫḫweni* ‘in the fire’ (dat),
  (c) PNWC: */-pʰa-/ ‘down,’ ‘to descend,’ WCirc /-pʰa-λaλa-/ -down-dangle-, Ub /-fa-/ ‘to ignite,’ */-x̂ʷə-/ ‘to fall,’ WCirc /-fə-/, ECirc /-x̂ʷə-/.

(65) ‘period of time,’ ‘season,’ ‘day’
  (a) Pontic: */məšə-(w)/ interval-predicative

(b) PIE: */məxʷə-r/ season-abs, Hitt *meḫur* ‘day, season,’ with Circassian-like development of */-š-w-/ > */-x̂ʷ-/,
*/məxʷə-la/ time-instr, Goth *mēl* ‘day’
*/məxʷə-ta-/ time-gen, Lat *mḗtior* ‘to measure out’
(c) PNWC: */məšə/ > PA-A */mšə/ ‘day’
*/məša-w/ time-predicative = ‘day’ > Kab /maax̂ʷa/, Ub /məx̂ʷá/ id (“məsʷá”)

(66) ‘sour, caustic liquid’
(a) Pontic */sax̂u/
(b) PIE: */səxʷə-r/ , Hitt *šeḫur* ‘urine,’ OIce *saurr* ‘semen,’ ‘impurity,’ ‘filth,’ *sṓggr*, *súrr* ‘sour,’ OE *sēaw*, Gk ὕει, Tokh B *sūwaṃ* ‘it rains’
(c) PNWC: */sax̂u/, Kab /sax̂ʷə/ ‘lime, quicklime’

(67) ‘people’
(a) Pontic: */rə-ĝa/ 3rd, impersonal-collective
(b) PIE: */a-rəĝa/ the-people > */haryo-/ , Hitt *arwa-* ‘free man’ (< **arya-wa-*, Ind-Iran **arya-* ‘Aryan,’ Gk ἀριστ- , Runic *arjostiz*, Welsh *irr* ‘charioteer,’ OIr *Airem* ‘a god’ (‘guardian of the Aryans’ ?) < **aryaman-*,
(c) PNWC: */(a-)rəĝa/ > Circ /adəĝa/, Ub /aː-dəĝa/ ‘Circassians,’ Abz /-rʕa/ ‘people’

(68) *‘house,’ ‘family’*
(a) Pontic: */ĝuna/ ‘house’
(b) PIE: */ĝuna-tʰa-qʰa/ > PIE */wuna-tʰ-qʰa/ house-of-belong, Dor Gk *ϝάναξ*, *ϝάνακτι* ‘lord’ (i.e., ‘head of the family’), Tokh A *nātăk*, Phrygian *ϝανακτει* id.
*/ĝuna-qʰa-ya-xa/ > PIE */wuna-qʰ-yaḥ/, Dor Gk *ϝάνασσα* ‘lady,’ Tokh A *nāśi* id.
(c) PNWC: */ĝuna/ > PCirc */wəna/ ‘house,’ Abz /ʕʷna/
*/ĝuna-tʰa-/ > PA-A */ĝʷəna-ta-ĝʷa/ house-gen-person = ‘family’ > Abz /ʕʷnatʕʷa/

(69) ‘man’
- (a) Pontic: */wə-ĝə-/male class marker-man- = ‘man’
- (b) PIE: */wə-ĝə-/ > PIE */wəy-/, Lat. *vīr*, Ir *fér*, Goth *wair*, Lith *vyras* ‘man,’ Sk *vayas* ‘strength’
- (c) PNWC: */wə-ĝə-/ > PC */ĝʷə/ > WCirc /(šʰə-)wə/ ‘(horse-)man,’ Ub /wə(də́)/ ‘devil’
  */wə-ĝə-/ > PA-A */ĝʷə/ > */ɣʷə/ > */ʕʷə/ > Abz /aʕʷə/ ‘man,’ /-ʕʷ/ ‘agent’
  */wə-ĝə-a/ > Ub /-ɣʷá/ sfx on pronouns

(70) ‘giant’
- (a) Pontic: */yən-ra/ gigantic-gerund = ‘the one who is big’
- (b) PIE: */yən-ra/ , Sk *Indra* (hero of the Rig Veda), Av *indra* ‘a demon,’ Hitt *innara* ‘a goddess’ (odd semantics of the PIE term are explained by Pontic)
- (c) PNWC: */yən(-ra)/, Circ /yənə/ ‘big,’ /yənə-ź/ big-evil = ‘giant,’ Abx /a-ynar/ ‘the-giant’

(71) ‘to say’
- (a) Pontic: */-(wə-)q’a-/ -(valence-)say- = ‘to say’ (‘to talk’)
- (b) PIE: */wə-q’a-/ > PIE */wəʔ-qʰa-/ -talk-belonging-(?) > */ʔəw-qʰʷ-/, Av *aok-* ‘to speak,’
  */ʔw-əqʰʷ-/, Ved *ví-vak-ti*, *Vakṣ*, Lat *uōx*, Umb *vepurus*, Gk (Ϝ)έπος
- (c) PNWC: */-(wə-)q’a-/ > WCirc /-ʔʷa-/, Kab /-ʔa-/, Ub /-q’a-/, Abx-Abz /-ḥʷa-/ ‘to say’

(72) ‘mouth’
- (a) Pontic: */ʕə-čʰa-/ edge-mouth = ‘lips,’ ‘mouth opening’
- (b) PIE: */a-wə-ʕə-čʰa-/ the-male-edge-mouth > PIE */haʕʷ-s-/, */hawʕ-s-/, Hitt *aiš-*, *išša-* (obl), Luw *āš*, Lat *ōs-*, *aus-*, Sk *ās-*, *óṣṫha-*,
- (c) PNWC: */wə-ʕə/ > PC */ʔʷə/ ‘mouth,’ ‘lips,’ ‘edge,’

*/ʕə-čʰa/ > P-Ub */ʕč̯a/ > */č̯'a/ > /(fa-)č'a/ '(nose-)mouth' = 'face'
*/ʕə-čʰə/ > PA-A */ʕč̯ə/ > */č̯'ə/ > Abx /(a-)č'ə́/ '(the-)mouth,'
*/yə-ʕə-čʰa/ > P-Ub */ʕč̯'a/ > */č̯'a/ > /č'a/ 'mouth'

(73) 'cattle'
  (a) Pontic: */wə-ʕə-(wə-ya-)/ male-cow/cattle-(being-one of-) = 'a grazing animal,'
  (b) PIE: */wə-ʕə-(wə-ya-)/ > PIE */ʕʷəw-y-/, Hitt *ḫawiš* 'sheep,' Luwian *ḫawi-*, Hieroglyphic Luwian *hawis*, Lat *ovis*, E *ewe*, Arm *hoviw* 'shepherd,'
  (c) PNWC: */wə-ʕə/, Circ /ʔʷəsə/ 'food, feed,' */wə-ʕə-a/, Circ /ʔʷa/ 'cattle pen'

(74) 'to be,' 'to be well'
  (a) Pontic: */ʔəčə-/ 'to be,'
  (b) PIE: */ʔəčə-/ > PIE */ʔəs-/ 'to be,' Sk *ás-thi*, Lat *est*, Goth *ist*, */ʔs-əw/ be-Adv = 'good, well,' Gk εὐ-, Sk *su-* (with lengthening of preceding vowels)
  (c) PNWC: */ʔəča-/ > */č'a/ > Ub /ča/ 'good,' by influence of the preverb form, */-ʔəča-wə-/ > P-Ub */-ć'ʷa-/ > /-śʷa-q'a-/ -good-say- = 'to speak well of someone,'
  */ʔəčə-wə-/ > PC */ć'ʷə-/, WCirc /ś'ʷə/ 'good,' Kab /f'ə/ id.

(75) 'two'
  (a) Pontic: */t'q'o/
  (b) PIE: */t'q'o/ > PIE */t'ʔʷə/, */t'əʔʷ/ > */dwō/, */do/ with levelling to */dwo/, Sk *dvā́*, *dvau*, OCS *duva*, Gk δύω, δύο, E *two*,
  (c) PNWC: */t'q'o/ > PC */t'ʔʷə/, P-Ub */t'q'ʷə/ > /t'q'ʷa/ orig. 'twice,' PA-A */t'ʕʷə/ > Abz /-ʕʷ/, Bzyb Abx /-yʷ/.

(76) 'six'
  (a) Pontic: */(w-)səx̂cə/ (masc class marker-)six,

(b) PIE: */(w-)səx̂cə/ > PIE */sʷəkʰs/, Gk ʽϝεξ (< */swəkʰs/), Lat *sex*, Goth *saihs* ,(both < */səkʰs/), Arm *veç*(< */wəkʰs/), OPruss *uschts* ‘sixth’ (< */wkʰs-tʰo-/), Av *xšwaš* < **švaš* (cf. *xšnaiti* < **zĭna:ti* ‘he knows,’ Gk γνώτι, E *know*), but perhaps by metathesis < */šwaxš/ < */sweks/,

(c) PNWC: */səx̂cə/ > PNWC */(s)x̂cə/ > PA-A */xcə/ > Abz /c-/, PC */x̂cə/ > */x̂sə/ > Circ /x̂ə/,
*/(w-)səx̂cə/ > P-Ub */x̂ʷcʷə/ > */śʷcʷə/ > Ub /fə/, PA-A */xʷċʷə/ > */śʷċʷə/ > Abx /f-/.

(77) ‘(hard) metal’

(a) Pontic: */(w-)ɣə-(č’a)/ ‘(grammatical class marker (I)-) metal-(hard),’

(b) PIE: */a-ɣʷc’ʷa/ > */hawc’ʷa/ > *ə$_4$*awso-* > Lat *aurum* ‘gold,’ */a-ɣʷc’ʷa/ > */haɣʷc’ʷa/ > *ə$_4$*e*ə$_3$*so-* > Lat *ōrum* id.

(c) PNWC: */ɣʷə-(č’a)/ > Bzh WCirc /ɣʷə-č’a/ ‘hard-metal’ = ‘iron,’ /ɣʷa-a-pƛa/ ‘metal-conn-red’ = ‘copper,’ Ub /wəc’ʷá/ ‘iron,’ Abz /ʕʷa(-t’a)/ ‘copper.’

(78) ‘metal (object)’

(a) Pontic: */yəx̂a/,

(b) PIE: */a-yx̂a/ > */hayx̂a/ > */hayšʸa/ > *ə$_4$*ay-so-*, *ə$_4$*y-es-* > Lat *aes*, Sk *áyas-* ‘metal,’ Av *ayah-* ‘metal object,’ Goth *aiz* ‘metal, money,’

(c) PNWC: */a-yəx̂a/ > Abx /a-ayxa/, Abz /ayxa/ ‘iron,’ ‘metal.’

(79) ‘son, child, foster child’

(a) Pontic: */pa/,

(b) PIE: */pa-w-/ > Gk *πάϝιδος > παῖς ‘child,’ παῦρος ‘little,’ Lat *puer* ‘boy,’ Sk *putra* ‘son,’ Osc *puklúm*, Paelignian *puclois*, Goth *fawai* ‘few,’

(c) PNWC: */pa-w-ʕə́-/ > PC */-paʔʷə́-/ > Bzh WCirc /-p’ʔʷə-/ ‘to rear,’
*/pa-w-ʕá-/ > PC */-paʔʷá-/ > Bzh WCirc /p’ʔʷa/ ‘foster

child,'
*/pa-y-ʕə-/ > Ub (northerly) /-p'q'ʸ-/ 'to rear,' */pa-ʕə-/ > Ub (southerly) */paɣ̄ə́/ > */pɣ̄ə́/ > /(q̄ʷa)p̄x̄ə́/ 'foster child,' */pa-ʕá/ > PA-A */px–á/ > /(qʷə-)px–á/ > Bzyb Abx /(a-)x–ʷpḥá/ 'foster child.'

(80) 'son,' 'nephew'
    (a) Pontic: */(nə-)pa-(tʰ-)/ '(lower-)son-(being/standing)' = 'nephew,
    (b) PIE: */nəpatʰ-/ > Lat. *nepōs*, Rumanian *nepot*, Ir *niae*, OE *nefa*, OHG *nevo*,
    (c) PNWC: */pa/ 'son.'

(81) 'to sit (down)'
    (a) Pontic: */(ʔa-)sə-(t'a-)/ '(change of state-)sit-(down-)'
    (b) PIE: */ʔəs-/ > Gk ἥ-μαι, ἥ-σ-ται, Hitt *e-eš-zi*, Sk *āste*, */ʔs-ət'-/ > Lat *sedēre*, Ir *saidim*, Lith *sėdėti*, Sk *sad-*, Goth *sitan*,
    (c) PNWC: */(ʔa-)sə-(t'a-)/ > Bzh WCirc /-qa-sə-ta-/ '-change of state-sit-down-' = 'to sit down' (with deglottalization of affixes), Ub /-s-/ 'to sit,' 'be situated' as in /a-s-q'a-y-á-s/ it-my-hand-dir-dat-sit = 'it is in my hand' (Vogt 1963:167, ⁄1457), /-t'ʷa-s-/ '-down-sit-' = 'to sit (down),' (with preposing of affix).

(82) 'to lie down, to fall down'
    (a) Pontic: */-ƛə-(g-y--)/ '-lie-(on-dat.)' = (1) 'to lie on,' (2) 'to fall on,'
    (b) PIE: */-ləgʸ-/ > Hitt *laki* 'causes to fall,' *lagāri* 'falls (mid.),' Gk λέχ-ομαι, Hom Gk λέκ-το, Lat *lectus* 'bed,' Ir *laigim*, Goth *ligan*, OCS *ležati*,
    (c) PNWC: */-ƛ-/ > PC */-ƛ-/ 'to lie, be prone,' Bzh WCirc /s-a-ƛ/ I-pres-lie = 'I am lying down,' for */-gʸə-/ note Ub /-gʸə-/ 'on' (preV);

*/-ƛ-a-/ -fall-dat- > PC */-ƛa-/ > Bzh WCirc /s-y-a-ƛa-a-ɣ/ I-dir-dat-fall-th v-past = 'I fell down'; with the same split in meaning as seen in PIE.

(83) 'sister'[6]
  (a) Pontic: */(w-)śə́mc'a/ '(class(I)-)woman,,'
  (b) PIE: */swə́s-ar/ 'woman-kin afx' > Sk *svasar-*, Lat. *soror*, Ir *siur*, Goth *swistar*, OCS *sestra*,
  (c) PNWC: */(w-)śə́mc'a/ > Ub /s^{w}ə́mc'a/ 'woman,' Bzyb Abx /(ḥ^w)ṡṡa/, WCirc /ś^{w}əz/, /pśaaśa/ 'girl' < */p-śəmc'a/ 'child-woman.'

Conclusions. First, PIE and PNWC are remotely related at a time depth of roughly 10,000 years.

Second, the sound system for the parent, Proto-Pontic, is likely that in (84).

(84) Proto-Pontic

p^h	p	b	-			m			w
t^h	t	d	t'			n	r	l	
c^h	c	ʒ	c'	s	z				
č^h	č	ǯ	č'	š	ž				y
ƛ^h	ƛ	λ	ƛ'						
k^h	k	g	k'	x̂	ĝ				
q^h	q	-	q'	x	ɣ				
			ḥ	ʕ					
		ʔ	h						
i	u								
e	ə,	o							
	a								

More work will have to be done to confirm all the vowels. The voiceless unaspirated series of stops is motivated by PNWC and seems to have fallen in with the voiceless aspirated stops in PIE. It is possible that this early loss led to later shifts and renewals in the source features of the voiceless stops in the various branches of Indo-European. Much more work is needed to trace out more complex sound laws. For example there are some sets where a labial-lateral cluster in NWC seems to correspond to a labiovelar in PIE, such as Circ /p'ƛ'ə/, Ub /p'ƛ'a/, A-A /pš^{y}/ all 'four'(which behaves as though it were a single segment in A-A, violating as it does the PA-A cluster rule *$C_1C_2 > C_2$), compared with PIE _*k^{w}etwer_ (*/k^{hw}ət^{h}wər/, or */k^{hw}ət^{hw}ər/) 'four.' It would seem from this vantage point that PIE was a gross simplification of Proto-Pontic. The history of the velar, uvular, pharyngeal, and laryngeal spirants, and /ʔ/ has already been delineated in (6)-(13). The affricates and spirants all seem to have fallen together into */s/, though further work is likely to show this to be an artifact of an overly simple image of PIE. The laterals seem all to have gone to */l/ , though here too further work is likely to yield interesting results.

Third, with its grammatical class prefixes (Colarusso 1989a) Proto-Pontic looks very much like a Daghestan or Northeast Caucasian language, and in fact further work is bound to show that PIE shares a phyletic link with PNEC as well, probably through Proto-North Caucasian, and perhaps with Proto-Kartvelian as well (Harris 1990).

Fourth, despite its NEC-look, PIE was spoken contiguously to PNWC, with some forms of PIE sharing some isoglosses with the more northerly portion (Proto-Circassian) of PNWC.

Fifth, the PIE homeland was most likely along the northeast shore of the Black Sea extending partially into the northwest region of the Caucasus, where its phyletic cousin dwelt. Proto-Pontic itself was likely to have been in the northwest Caucasus, extending up into what is now the Crimea and southern Ukraine. The steppe offered opportunities to exploit the horse in a nomadic economy, and this opportunity set the ancestors of the Indo-Europeans apart from their kinsmen in the mountains and launched them upon the stage of history.

NOTES

[1] The amateur archaeologist, Geoffry Bibby, suggested in 1961 that PIE was a Caucasic language that went north and blended with a Finno-Ugric tongue. This guess seems to owe more to the old notion that the Caucasus was the source for many of the peoples of Europe than it does to an informed notion of PIE, of any Caucasic languages, or of Finno-Ugrian. Friedrich's conjecture, therefore, takes historical precedent.

[2] I use 'Caucasic' rather than the more traditional 'Caucasian' to avoid any naive confusion that somehow these are "white man's languages."

[3] Given some of the recent publicity (Ross 1991, Wright 1991) surrounding the revival of the late nineteenth-century notion that every language is ultimately related to every other (Pedersen 1931:338-339), I wish explicitly to dissociate myself from any such efforts. In fact most such notions try to link North Caucasic languages with those in Asia, such as Sino-Tibetan or Yeniseian, or even more remotely with the Amerindian Na-Dené, while linking PIE with Uralo-Yukaghir, South Caucasian (Kartvelian) or Elamo-Dravidian and Afro-Asiatic (Ross: 138-139). The plausibility of what follows simply shows the folly of such grand lumping schemes.

[4] There is one Northeast Caucasian language, the Richa dialect of Aghul, which actually contrasts these types of sounds (Kodzasov 1987). In the back of the mouth it contrasts uvulars: pharyngealized uvulars: pharyngeals: adytals: V̌ = a pharyngealized V: /xă/ 'house,' /x̄ăw/ 'nut,' /h̯aw/ 'udder,' /ḥač/ 'apple'; /ɣad/ 'hammer,' /ɣ̄ăb/ 'stack,' /ʕ̯an/ 'belly,' /ʕakʷ/ 'light' [my re-transcription].

[5] There are a number of resemblances between PIE and Proto-Kartvelian (Howard Aronson, personal communication; Alice Harris 1990; Gamkrelidze and Ivanov 1967; Gamkrelidze 1966), so much so that an investigation similar to this one is warranted. Phyletic links between PIE and Proto-Kartvelian would, of course, establish PIE as an outlier of an ancient Proto-Caucasic.

[6] Eric Hamp (personal comunication) has suggested that the root here is merely */sar-/, with */swə-/ being the reflexive. His argument is based upon the Latin pair *soror* (< */swəsar-/) vs. *uxor* 'wife.' This has a parallel in Vajzë Albanian *r-ya-* 'woman-diminutive-' = 'wife' vs. *var-ya-* 'sister-diminutive-' with *v-ar-* < */swə-sar-/. If the Albanian form is not a parallel built upon Latin influence but rather derived from Indo-European patterns, then it would suggest that the PIE was */swə-sar-/ 'own-woman' = 'sister,' */ukʰ-sar-/ 'outer-woman' = 'wife,' and this Pontic match would have to be rejected.

REFERENCES

Abdokov, A. I. 1983. O zvukovyx i slovarnyx sootvetstvijax severokavkazskix jazykov. Nal'čik: El'bruz.

Allen, W. Sidney. 1965. "On One Vowel Systems." Lingua 13:111-124.

Benveniste, Émile. 1935. Origines de la formation des noms en indo-européen. Paris: Adrien-Maissonneuve. [1962 reprint].

Bibby, Geoffrey. 1961. Four Thousand Years Ago. New York: Alfred A. Knopf.

Brugmann, Karl. 1888. Elements of the Comparative Grammar of the Indo-Germanic Languages. Joseph Wright (trans.). Strassburg and London: Trübner & Co.

Buck, Carl Darling. 1949. A Dictionary of Selected Synonyms in the Principal Indo-European Languages. Chicago: University of Chicago Press.

Čirikba, Vjaččeslav Andrejevič. 1986. Sistema svistjaščix soglasnýx v abxazo-adygskix jazykax. Moscow: Institut jazykoznanija AN SSSR.

Colarusso, John 1981. Typological Parallels between Proto-Indo-European and the Northwest Caucasian Languages. In Yoël Arbeitman and Allan R. Bomhard (eds.), Bono Homini Donum: Essays in Historical Linguistics in Memory of J. Alexander Kerns , vol. 2, pp. 475-558. Amsterdam: John Benjamins.

_____. 1984. Parallels between the Circassian Nart Sagas, the *Rg Veda*, and Germanic Mythology, in V. Setty Pendakur (ed.), South Asian Horizons, vol. 1, Culture and Philosophy, pp. 1-28. Ottawa: Carleton University, Canadian Asian Studies Association.

_____. 1985. Pharyngeals and Pharyngealization. IJAL 51.4: 366-368.

_____. 1989a. Proto-Northwest Caucasian, or How to Crack a Very Hard Nut. In Howard I. Aronson (ed.),The Non-Slavic Languages of the USSR, Linguistic Studies, University of Chicago: Chicago Linguistic Society, pp. 20-55.

_____. 1989b. The Woman of the Myths: the Satanaya Cycle, in Howard I. Aronson (ed.), The Annual of the Society for the Study of Caucasia 2: 3-11.

Diakonoff, Igor M. 1990. Language Contacts in the Caucasus and the Near East. In T. L. Markey and John A. C. Greppin (eds.) When Worlds Collide; Indo-Europeans and Pre-Indo-Europeans. Ann Arbor, Michigan: Karoma Publishers, Inc. Pp. 53-65.

Friedrich, Paul. 1964. Review of *Phoneme and Morpheme in Kabardian (Eastern Adyghe),* Aert Kuipers. (Janua Linguarum. Studia Memoriae Nicolai Van Wijk Dedicata, No. VIII). The Hague: Mouton and Co., 1960. 124 pp., appendix, bibliography, tables. f. 16. American Anthropologist 66:205-209.

Gamkrelidze, Thomas V. 1966. A Typology of Common Kartvealian. Language 42:69-83.

Gamkrelidze, Thomas V. and Ivanov, V. V. 1967. Kartvelian and Indo-European, a Typological Comparison of Reconstructed Systems. In To Honor Roman Jakobson, vol. 1, pp. 707-717 The Hague: Mouton.

_____. 1972. Lingvističeskaja tipologija i rekonstrukcija sistemu indoevropejskix smyčnyx. Working Papers of the Conference on the Comparative-Historical Grammar of the Indo-European Languages (12-14 December 1972). Moscow. Pp. 15-18.

_____. 1973. Sprachtypologie und die Rekosntruktion der gemeinindogermanischen Verschlüsse. Phonetica 27:150-156.

_____. 1984. Indoevropejskije jazyki i indojevropejcy. Tbilisi: Tbilisi University Press.

_____. 1985. The Ancient Near East and the Indo-European Question [and] the Migration of Tribes Speaking Indo-European Dialects. JIES 13:3-91.

Gamqrelije [Gamkrelidze], Tamaz and Mač'avariani, Givi. 1965. Sonant'ta sist'ema da ablaut'i kartvelur enebši [The Sonant System and Ablaut in the Kartvelian Languages]. (In Georgian with Russian summary) Tbilisi.

Gimbutas, Marija 1973. The Beginning of the Bronze Age in Europe and the Indo-Europeans: 3500-2500 B.C. JIES 1: 163-214.

_____. 1974. An Archaeologist's View of PIE in 1975. JIES 2:289-308.

_____. 1977. The First Wave of Eurasian Steppe Pastoralists into Copper Age Europe. JIES 5:277-338.

_____. 1980. The Kurgan Wave 2 (c. 3400-3200 BC) into Europe and the Following Transformation of Culture. JIES 8:273-315.

_____. 1985. Primary and Secondary Homeland of the Indo-Europeans. JIES 13:185-202.

Goddard, Ives 1975. Algonquian, Wiyot and Yurok: Proving a Distant Genetic Relationship. In M. Dale Kinkade, Kenneth L. Hale and Oswald Werner (eds.), Linguistics and Anthropology: In Honor of C. F. Voegelin, pp. 249-262. Lisse: The Peter de Ridder Press.

Hamp, Eric P. 1990. The Indo-European Horse. In T. L. Markey and John A. C. Greppin (eds.) When Worlds Collide; Indo-Europeans and Pre-Indo-Europeans. Ann Arbor, Michigan: Karoma Publishers, Inc. Pp. 211-226.

Harris, Alice C. 1990. Kartvelian Contacts with Indo-European. In T. L. Markey and John A. C. Greppin (eds.) When Worlds Collide; Indo-Europeans and Pre-Indo-Europeans. Ann Arbor, Michigan: Karoma Publishers, Inc. Pp. 67-100.

Hopper, Paul J. 1973. Glottalized and Murmured Occlusives in Indo-European. Glossa 7:141-166.

______. 1977a. The Typology of the Proto-Indo-European Segmental Inventory. JIES 5:41-54.

______. 1977b. Indo-European Consonantism and the New Look. Orbis 26:57-72.

______. 1982. Areal Tupology and the Eraly Indo-European Consonant System. In Edgar C. Polomé (ed.) The Indo-Europeans in the Fourth and Third Millenia. Ann Arbor, Michigan: Karoma Publishers. Pp. 121-139.

Jasanoff, Jay 1978. Stative and Middle in Indo-European. Innsbrucker Beiträge zur Sprachwissenschaft.

Kodzasov, Sergei V. 1987. Pharyngeal Features in the Daghestan Languages. Proceedings of the XIth International Congress of Phonetic Sciences, vol. 2, pp. 142-144. Tallinn, Estonia.

Kuipers, Aert H. 1960. Phoneme and Morpheme in Kabradian. The Hague: Mouton.

______. 1975. A Dictionary of Proto-Circassian Roots. Louvain, Belgium: Peeters.

______. 1983. Review. Thomas V. Gamkrelidze and Givi I. Mačavariani, Sonantensystem und Ablaut in den Kartwelsprachen. Eine Typologie der Struktur des Gemeinkartwelischen. Mit einem Vorwort von Georg Tsereteli, Ins Deutsche übersetzt, bearbeitet und mit einem Nachwort von Winfred Boeder. [Ars Linguistica 10. Commentationes analyticæ et criticæ]. Tübingen: Gunter Narr Verlag 1982. [160 pp.]. Studia Caucasica 5:98-102.

Kuryłowicz, Jerzy 1964. The Inflectional Categories of Indo-European. Heidelberg: Carl Winter.

Lehmann, Winfred P. 1952. Proto-Indo-European Phonology. Austin: University of Texas Press.

Lindeman, Fredrik Otto 1990. Is There Any Conclusive Evidence for a Triple Representation of Schwa in Armenian? Annual of Armenian Linguistics 11:25-30.

______. 1987. Introduction to the 'Laryngeal Theory.' Oslo: the Norwegian University Press, the Institute for Comparative Research in Human Culture.

Mallory, J. P. 1989. In Search of the Indo-Europeans, Language Archaeology and Myth. London: Thames & Hudson.

Martinet, André 1986. Des steppes aux océans, L'indo-européen et les "Indo-Européens." Paris: Payot.

Meillet, Antoine 1922 [1964 printing]. Introduction à l'étude comparative des languages indo-européennes. University of Alabama Press.

Pedersen, Holger 1931. The Discovery of Language. Translated by John Webster Spargo. Bloomington: Indiana University Press, 1962 edition.

Pisani, Vittore 1947. Crestomazia indeuropea. Torino: Rosenberg & Sellier.

Ross, Philip E. 1991. Hard Words. Scientific American, vol. 264, no. 4, April, pp. 138-147.

Vogt, Hans 1963. Dictionnaire de la langue oubykh. Oslo: Universitetsforlaget.

Watkins, Calvert 1980. Indo-European and the Indo-Europeans; Guide to the Appendix; Indo-European Roots. In The Houghton-Mifflin Canadian Dictionary of the English Language. Pp. 1496-1550.

Winter, Werner (ed.) 1965. Evidence for Laryngeals. The Hague: Mouton.

______. 1970. Some Widespread Indo-European Titles. In George Cardona, Henry M. Hoenigswald and Alfred Senn (eds.) Indo-European and the Indo-Europeans. Philadelphia: University of Pennsylvania Press. Pp. 49-54.

Wright, Robert 1991. Quest for the Mother Tongue. The Atlantic, vol. 267, no. 4, April, pp. 39-68.

FOCUS IN YUKAGHIR (TUNDRA DIALECT)

Bernard Comrie
University of Southern California

ABSTRACT

A number of extensions and further generalizations are made to Krejnovič's (1958) account of focus in Yukaghir. Krejnovič distinguishes subject, object, and predicate focus; a fourth type, neutral focus, must be recognized, in particular where some element other than one of these three (e.g., an adverbial) is in focus. Interrogative words are necessarily in focus. Focus oppositions are neutralized in nonfinite subordinate, imperative, and negative sentences. These further generalizations enable Yukaghir focus to be integrated more fully into crosslinguistic studies of grammaticalized focus.

One of the characteristics of the Yukaghir language, as described by Krejnovič (1958), is the existence of a rich system of morphological means for the encoding of the focus of a sentence, where 'focus' is understood as the essential new information conveyed by the sentence.[1] The aim of this article is to systematize and elaborate Krejnovič's discussion of this phenomenon. The material on which the discussion is based is restricted to the material presented in Krejnovič (1958), more specifically to his material on the Tundra dialect of Yukaghir, including both illustrative paradigms and sentences from the body of the book and examples from the text on pages 255-262. Examples from Krejnovič (1958) are identified either by page number (for sentences from the body of the book) or by the abbreviation T followed by the sentence number (for sentences from the text, using Krejnovič's numbering). My indebtedness to the late E.A. Krejnovič's work will be evident at every turn, and I hope that this article may stand as a small token of appreciation for his work.

1. The basic system. According to Krejnovič, clauses may appear in three variants in Yukaghir: predicate focus, subject focus, and [direct] object focus; intransitive clauses, of course, may appear only in the first two variants. Focus is shown by a rather complex interplay of verbal and nominal morphology; for the verbal morphology, see the appendix. The three focus variants will be treated below in the order just given.

In predicate focus (Krejnovič 1958:131-133, 146-148), the verb takes the prefix *me(r)-* and is conjugated for person-number, with different person-number suffixes for intransitive and transitive verbs. Subjects, whether transitive or intransitive, take no case ending (here called absolute case). Direct objects also usually stand in the absolute case, except in the following circumstances: (a.) a direct object noun takes the suffix *-le* if the subject is also third person (Krejnovič 1958:47-48); (b.) first and second person pronouns have a special accusative case in *-ul* if the subject is also first or second person; (c.) if, however, the subject is third person, then first and second person direct objects stand in the locative (Krejnovič 1958:75). The various possibilities are illustrated below.

(1) met mer-ū-jeŋ. (146)
I FOC-go-1SG
'I *went.*'

(2) kideŋ mer-ū-j. (146)
person FOC-go-3SG
'the person *went.*'

(3) met nimeŋ me-wie-ŋ. (47)
I house FOC-make-1SG
'I *built* a house.'

(4) tudel nime-le me-wie-m. (48)
he house-ACC FOC-make-3SG
'he *built* a house.'

(5) met tet-ul me-paj-ŋ. (75)
I you-ACC FOC-hit-1SG.
'I *hit* you.'

(6) tudel met-qane me-pótesej-m. (76)
he I-LOC FOC-send-3SG

'he *sent* me.'

Under subject focus (Krejnovič 1958:133-135, 152-154), the verb stands in a nonfinite form, i.e., one lacking person-number agreement (though for third person plural there is a number suffix *-ŋu);* for transitive verbs, this nonfinite form is simply the stem, while for intransitive verbs a suffix *-l* is added. In intransitive clauses, the subject takes the suffix *-le(ŋ)* (nouns without attributes) or *-(e)k* (nouns with attributes and first and second person pronouns); a few noun phrases, including third person pronouns, remain unmarked. In transitive clauses, the subject remains unmarked (Krejnovič 1958:37-45). Various possibilities are illustrated below.

(7) met-ek ū-l. (152)
I-FOC go-NONFIN
'*I* went.'

(8) tittel ū-ŋu-l. (152)
they go-PL-NONFIN
'*they* went.'

(9) ile-leŋ ū-l. (152)
reindeer-FOC go-NONFIN
'*the reindeer* went.'

(10) met (ileŋ) ai.[2] (133)
I reindeer shoot:NONFIN
'*I* shot the reindeer.'

Under object focus (Krejnovič 1958:135-138), the object takes the same form as a focused intransitive subject (Krejnovič 1958:37-45), while the verb agrees in person-number with the subject, though using inflections somewhat different from those used under predicate focus (intransitive or transitive), as illustrated below.

(11) met ile-leŋ ai-meŋ. (135)
I reindeer-FOC shoot-1SG(OBJ:FOC)
'I shot *the reindeer.*'

It will be noted that the nominal focus marker is distributed on an ergative-absolutive basis, occurring with intransitive subjects and transitive direct objects, but not with transitive subjects. This seems, incidentally, to be a fairly isolated instance of ergativity in Yukaghir; note, for instance, that the use of nonfinite verb forms under subject focus is the same for both focused intransitive and focused transitive subjects.

A number of questions arise on the basis of the brief description above, and the aim of the remainder of this article is to try and answer some of these questions. One question concerns possible restrictions on the occurrence of the focusing variants: is the full paradigm possible in all kinds of clauses? While Krejnovič discusses explicitly focus on the predicate, the subject, and the direct object, what of other kinds of focus, for instance on various oblique objects or on adverbials? Krejnovič's presentation essentially assumes that each clause (or at least each clause for which focus is relevant) must occur in one of the three focus forms presented above; but is there some way of indicating absence of (explicit) focus, or is one of these three focus types pragmatically neutral, i.e., used when it is not desired to focus any particular constituent?

In fact, in addition to the three focus types noted explicitly by Krejnovič, there is a fourth type, mentioned in passing (Krejnovič 1958:151-152), identical to predicate focus except that the verb lacks the focus prefix *me(r)-*. I will call this (formal) neutral focus, since none of predicate, subject, or direct object (or, indeed, any other constituent of the sentence) is marked as focus; examples of neutral focus are (12), (13) (in the subordinate clause only), and (14) (contrast (15), with subject focus).

(12) id́ilwej-mirije lat́il-ŋiń kewe-t́. (150)

PN-woman firewood-DIR go-3SG
‘Ídilwej went for firẹwood/*firewood.*’[3]

(13) t́ārt́eqan lalime tudie-reŋ me-kewe-t́. (150)
PN sledge drag-SIMUL:GER FOC-go-3SG
“T́ārt́eqan *went*, dragging the sledge.”[4]

(14) tenin, aweja, t́īŋ ū́ti-ŋi. (206)
here yesterday people pass-3PL
‘here, yesterday, people passed by.’[5]

(15) tenin idíre, jaŋre-pu-leŋ ū́ti-ŋu-l. (206)
here now goose-PL-FOC pass-PL-NONFIN
‘here, just now, *geese* passed by.’

Examples of neutral focus will play a major role in the following discussion.

2. Interrogative pronouns and adverbials. To start this section, let us consider interrogative pronouns appearing in subject or direct object position. Krejnovič gives a rather complex set of rules for the form of subject and direct object interrogative pronouns, which I will try to simplify considerably. According to Krejnovič (1958:88), the subject form for ‘who’ is *kin* with intransitive verbs, and *kinek* with transitive verbs; he gives no examples of ‘who’ as direct object of a transitive verb. According to Krejnovič, ‘what’ appears as *nemeŋ* or *neme-leŋ*, the latter as intransitive subject with subject focus and as direct object with object focus; he gives no examples of ‘what’ as subject of a transitive verb. The rules given by Krejnovič can be replaced by the following simple generalization: interrogative pronouns must always be the focus of their sentence. It thus follows that the morphology of the interrogative pronoun and of the verb will be precisely that predicted by having the focused element in the grammatical relation of the interrogative pronoun. This is in keeping with the function of interrogative pronouns (to request essential new information), and also with their crosslinguistic behavior in languages with grammaticalized focus; see, for instance, Kim (1988:151). Illustrative examples follow.

(16) kin-ek kelu-l. (88)
who-FOC come-NONFIN
'*who* came?'

(17) kin ket́i. (88)
who bring:NONFIN
'*who* brought [it]?'

(18) *neme-le quduole-l.* (88)
what-FOC lie-NONFIN
'*what* lies [here]?'

(19) *neme-le wie-te-meŋ.* (88)
what-FOC do-FUT-2SG(OBJ:FOC)
'*what* will you do?'

(20) *kədeŋ quodeband́-ile-k puń-mele.* (138)
person what-reindeer-FOC kill-3SG(OBJ:FOC)
'*what* kind of reindeer did the man kill?'[6]

(21) *kədeŋ amat́ed-ile-k puń-mele.* [answer to (21)] (138)
person good-reindeer-FOC kill-3SG(OBJ:FOC)
'the person killed a *good* reindeer.'

The examples given above of interrogative pronouns all involve interrogative pronouns as subject or direct object. What happens if the interrogative pronoun bears some other grammatical or semantic relation, such as comitative, locative? The following examples illustrate various such possibilities.

(22) kin-han pot́esej-mek. (88)
who-PROL send-2SG
'*with whom* did you send [it]?'

(23) neme-han tadi-mek? (89)
what-PROL give-2SG
'*in exchange for what* did you give [it]?'

(24) pohode-han tadi-ŋ. [answer to (24)] (89)
money-PROL give-1SG
'I gave it *in exchange for money.*'

(25) met-ul qodaŋ pun-te-mek. (205)
I-ACC how kill-FUT-2SG
'*how* will you kill me?'

(26) quodir kədeŋ qaŕtuora-nu-mek. (206)
why person deceive-IPFV-2SG
'*why* do you deceive people?'

(27) qadat kelu-k.[7] (207)
whence go-IMPER:2SG
'*from where* have you come?'

(28) nemol lew-te-m. (206)
why eat(TR)-FUT-3SG
'*why* will he eat it?'

In all of these examples, and indeed in all comparable examples given by Krejnovič, the verb is in neutral focus.[8] This observation fits in readily with several observations made above. If it is true that an interrogative pronoun is necessarily in focus position, then in examples where the interrogative pronoun bears some relation other than subject or direct object (or, of course, predicate), it follows that the construction can be neither subject focus, nor object focus, nor predicate focus. This, then, illustrates one environment where neutral focus is required.

But if adverbials expressed by interrogative pronouns stand in focus position, then one would expect that other kinds of adverbials, if functioning as focus of the sentence, should also appear in focus position. Krejnovič (1958:151) actually makes a stronger claim, namely that with an adverb of quality one must use neutral focus; however, examination of his own data shows that this generalization is too broad. Rather, as predicted by our general line of argument, adverbials appear in focus position if they are essential new information ((29)-(33)), and not otherwise ((34)-(35)). The following examples illustrate adverbials in and out of focus position. Since there is no overt positive indication of focus on an adverbial, it is also possible that some of the examples cited below with focus on the adverbial might conceivably rather have no focus, a point to which I return in section 4.

(29) kədeŋ amut-neŋ əlke-j. (151)
person good-ADV run-3SG
'the person ran *well.*'

(30) kin-ek ńārt́it-ne enuskurie-l, ńārt́it-ne lewde-j. (T318)
who-FOC bad-ADV work-NONFIN bad-ADV eat (ITR)-3SG
'*whoever* works badly, eats *badly.*' [9]

(31) t́amaneŋ it-neŋ ū-ŋi. (203)
very long-ADV go-3PL
'they walked *for a very long time.*'

(32) wien-lukunburube-hat teni ket́i-ŋa. (207)
other-land-ABL here bring-3PL
'they brought it here *from another land.*' [10]

(33) sa-l-ha welte-m, puren. (209)
tree-LOC hang-3SG above
'he hung it *on a tree*, high up.'

(34) id́ire mit-sahanel wari-neŋ t́ald́e-ha me-med́i-j. (202)
now we-life firm-ADV hand-LOC FOC-take-1PL
'now we *have taken* our life firmly in [our] hands.'

(35) tada me-panre-ŋ. (T54)
there FOC-set(trap)-1SG
'I *set* [a trap] there.'

On the basis of this discussion, it is now clear that sentence (12) should probably be analyzed as having 'firewood' as its focus.

3. Restrictions on the occurrence of focus oppositions.

3.1 Nonfinite dependent clauses. In Yukaghir, certain types of sentence do not show focus oppositions. The system outlined in sections 1-2 essentially characterizes only finite clauses, and in a Yukaghir sentence there is usually only one finite clause; the finite subordinate clauses of a language like English are typically expressed by means of nonfinite forms, in particular participles, gerunds, and verbal nouns, and none of these nonfinite clause types permits the expression of focus distinctions.[11] Case marking of noun phrases in clauses outside the focus system is as under predicate focus, but the predicate itself does not bear the predicate focus prefix *me(r)-* (Krejnovič 1958:151); the structure is thus parallel to neutral focus in finite clauses.

Sentence (13) above, repeated as (36), illustrates absence of focus marking in a dependent clause.

(36) t́āŕteqan lalime tudie-reŋ me-kewe-t́. (150)
PN sledge drag-SIMUL:GER FOC-go-3SG
'T́āŕteqan *went*, dragging the sledge.'

3.2 Imperative sentences. Certain types of finite clauses also lack focus oppositions. For instance, in imperative sentences, the focus system is suspended, the noun phrases being encoded as under predicate focus but the verb lacking the focus prefix *me(r)-* (Krejnovič 1958:150-151). It is frequent crosslinguistically for imperative verbs to be necessarily focused, thus preempting the expression of focus elsewhere.[12]

(37) wie-k. (127)
make-IMPER:2SG
'do [it]!'

(38) kelu-k. (129)
come-IMPER:2SG
'come!'

3.3 Negative sentences. Another example of loss of focus oppositions is provided by negative finite clauses (Krejnovič 1958:138-140, 151-152). These have the basic structure of predicate focus clauses, except that instead of the predicate focus prefix *me(r)-* we find the negative prefix *el-*. In many languages with grammaticalized marking of focus, the marker of negation preempts any other overt marking of focus, i.e., grammatically the negation marker is the focus.[13] This principle provides an elegant way of handling the behavior of Yukaghir focus under negation: the negative prefix is the focus, thus usurping the position of the predicate focus marker *me(r)-* and also preventing focus from being marked on any other constituent of the clause. The following parallel examples illustrate this phenomenon:

(39) met mer-ū-jeŋ. (151)
I FOC-go-1SG
'I *went.*'

(40) met eĺ-ū-jeŋ. (151)
I NEG-go-1SG
'I didn't (did *not)* go.'

With transitive verbs, Yukaghir uses intransitive conjugation under negation, so that the person-number inflections are not parallel in the next pair of sentences (Krejnovič 1958:138-140). Hopper and Thompson (1980:276-277) note that in many languages negative sentences behave as 'less transitive' than their positive counterparts.

(41) met ile me-puń-iŋ. (42)
I reindeer FOC-kill-1SG
'I *killed* the reindeer.'

(42) met ileŋ eĺ-buń-d́eŋ. (140)
I reindeer NEG-kill-1SG
'I didn't (did *not)* kill the reindeer.'

There is, however, one exception to the general pattern of negation. If the direct object is qualified by an attribute (and perhaps in a few other instances), the object stands in the focus form and the verb remains transitive (object-focus) in inflection (Krejnovič 1958:140).

(43) met amat́ed-ile-k eĺ-buń-meŋ. (140)
I good-reindeer-FOC NEG-kill-1SG(OBJ:FOC)
'I did not kill the *good* reindeer.'

While the precise details of this pattern seem somewhat idiosyncratic, perhaps meriting further investigation, note that the direct object in (43) is more individuated than in (42), therefore the overall construction more transitive, following the general lines of Hopper and Thompson (1980:252-253).

4. Are there truly focusless sentences in Yukaghir? If we return to sentence types where focus oppositions are expressible, then we may also return to a question posed in section 1, namely whether it is possible for a Yukaghir sentence of this type to have no focus, for instance because all information in the sentence is presented as new, or simply because the speaker does not want to make explicit what information is new in the sentence. In fact, once one notes that what I have been calling neutral focus means that the focus is neither subject, direct object, nor predicate, but may for instance be an adverbial, there are virtually no instances of truly focusless sentences in Krejnovič's materials. Indeed, one characteristic of Krejnovic's Yukaghir text is the prevalence of sentences where the verb has the prefix *me(r)-*, suggesting the conclusion that unless some constituent other than the predicate is marked as focus, then by default (the 'pragmatically neutral case') the predicate is marked as focus. Of the few examples that might be considered exceptions, one is (14), repeated below.

(44) tenin, aweja, t́īŋ ū́ti-ŋi. (206)
here yesterday people pass-3PL
'here, yesterday, people passed by.'

From the form of the verb and subject in (44) (neutral focus), it is clear that neither the verb nor the subject is focus. Grammatically, either the place adverbial or the time adverbial could be focus, but from the way in which they are presented, preposed and separated by commas, the suggestion is that they form topical, rather than focused, material. The resolution of the analysis of this particular example will of course require much more detailed work on Yukaghir than is possible using the limited textual material provided by Krejnovič. There is, however, another type of example which seems to have truly neutral focus, namely when the verb *mon* 'say' is used to introduce direct speech, as in the following example.

(45) mon-ŋi: qamlal-mut. (T407)
say-3PL: how:many-2PL
'they said: how many of you [are there]?'

Morphologically, the main clause of (45) is in neutral focus. Since there is no expressed subject in (45), it cannot be an instance of subject focus, and in any case the verb is not in the nonfinite form that would be required by subject focus. The verb lacks the prefix *me(r)-*, so is not in predicate focus. It is, moreover, clear from the morphology of the verb form *mon-ŋi* that this verb is intransitive (at least in this usage), so there can be no question of object focus. Finally, the main clause of (45) contains no other constituent that could be focus. Indeed, the only possibility of analyzing (45) as having a focused constituent is to suppose that the quotation *qamlal-mut* is the focus, i.e., that this quotation is a constituent of the main clause but not its direct object; yet this would seem to contradict the generalization that the expression of focus in Yukaghir is clause-bound (see section 3.1 and note 7).

5. Conclusions. In addition to the three focus types recognized explicitly by Krejnovič (subject, object, and predicate focus), it is necessary to distinguish a fourth type, neutral focus, which lacks any explicit indication of focus and occurs most readily when some constituent other than the subject, object, or predicate is focused. Formally, neutral focus is the simplest form, and as such one might wonder why Krejnovič does not treat it as basic; but given the tendency (perhaps an absolute rule) for some constituent to be marked as focus in a Yukaghir finite (affirmative, nonimperative) sentence, neutral focus is in fact rather rare.

Appendix: Yukaghir verbal morphology

(*ū-* 'go,' *ai-* 'shoot,' Aorist tense; Krejnovič 1958:131, 133, 135, 146, 152)

	Intransitive	Transitive
Predicate focus		
SG1	mer-ū-jeŋ	mer-ai-ŋ
2	mer-ū-jek	mer-ai-mek
3	mer-ū-j	mer-ai-m
PL1	mer-ū-jeli	mer-ai-j
2	mer-ū-jemut	mer-ai-mk
3	mer-ū-ŋi	mer-ai-ŋa
Subject focus		
Other	ū-l	ai
PL3	ū-ŋu-l	ai-ŋu
Object focus		
SG1		ai-meŋ
2		ai-meŋ
3		ai-mele
Pl1		ai-l
2		ai-mk
3		ai-ŋumle

REFERENCES

Comrie, Bernard. 1984. Some formal properties of focus in Modern Eastern Armenian. Annual of Armenian Linguistics 5.1-21.

Hopper, Paul J. and Thompson, Sandra A. 1980. Transitivity in grammar and discourse. Language 56.251-299.

Kim, Alan Hyun-Oak. 1988. Preverbal focusing and type XXIII languages. Studies in syntactic typology (Typological Studies in Language, 17), ed. by Michael Hammond, Edith Moravcsik, and Jessica Wirth, 147-169. Amsterdam: John Benjamins.

Krejnovič, E.A. 1958. Jukagirskij jazyk. Moscow--Leningrad: Izdatel'stvo AN SSSR.

Kuno, Susumu. 1973. The structure of the Japanese language. (Current Studies in Linguistics Series, 3.) Cambridge, MA: MIT Press.

NOTES

[1] A slightly earlier version of this paper was presented at the Sixth International Conference on the Non-Slavic Languages of the USSR. I am grateful to all those who contributed to the discussion following my presentation. The following abbreviations are used: ABL — ablative; ABS — absolute; ADV — adverb; DIR — directional; FOC — focus; FUT — future; IMPER — imperative; INST — instrumental; IPFV — imperfective; ITR — intransitive; LOC — locative; NEG — negative; NONFIN — nonfinite; OBJ:FOC — object focus; PL — plural; PN — proper name; PROL — prolative; SG — singular; SIMUL:GER — simultaneous gerund; TR — transitive. The focus of a sentence is indicated by italicizing the focused constituent in the English translation.

[2] As is illustrated by the version of example (10) including the parenthetical material, as well as by several other examples below, in Yukaghir there is no rule, certainly no absolute rule, that the focus must immediately precede the predicate, unlike a number of other basically verb-final languages with grammaticalized focus (Kim 1988). Prepredicate position cannot therefore be used as a criterion for identifying focus in Yukaghir.

[3] See the further discussion of this sentence in section 2.

[4] See the further discussion of this sentence in section 3.1.

[5] See the further discussion of this sentence in section 4.

[6] Note in this and the following example (and also (43) below) that the (pragmatic) focus is strictly speaking a subconstituent of the direct object. In Yukaghir, as usually in languages with grammaticalized focus, pragmatic focus on a subconstituent requires grammatical focus of the whole constituent.

[7] With certain interrogative adverbials, imperative second person verb forms are used rather than indicative verb forms (Krejnovič 1958:129).

[8] The following question and answer pair provide an unexpected exception to the general pattern. In the question, the interrogative pronoun, semantically an instrumental (judging from Krejnovič's Russian translation), is focus, and as expected the

construction is neutral focus. In the answer, however, the construction is unequivocally object focus, although the direct object is just an unexpressed 'it.' Either (ii) is an error (the verb should be *wie-te-ŋ)*, or object focus can sometimes be extended to non-subject arguments/adjuncts other than the direct object.

(i) tet neme-lek wie-te-mek? (88)
you what-INST do-FUT-2SG
'*with what* will you make it?'

(ii) nimud́i-lek wie-te-meŋ. [answer to (i)] (88)
ax-INST do-FUT-1SG(OBJ:FOC)
'I will make it *with an ax*.'

[9] Note that this correlative construction contains two finite clauses. The first has subject focus (as expected from the occurrence of an interrogative pronoun in subject position), the second neutral focus.

[10] The focus could also be *'here.'*

[11] Compare Kuno's (1963:56-57) account of the neutralization of thematic oppositions in subordinate clauses in Japanese.

[12] For instance, in Hungarian verbal prefixes normally occupy the preverbal focus position, e.g., *el-ment-él az iskolá-ba* 'away-go:PAST-2SG the school-DIR,' i.e., 'you went (away) to school,' with *el-* 'away' in focus position; in the imperative, however, the verbal prefix is normally placed after the verb, i.e., does not occupy the preverbal focus position, e.g., *menj el az iskolá-ba* 'go-IMPER:2SG away the school-DIR,' i.e., 'go (away) to school.'

[13] Compare Hungarian *nem ment-él el az iskolá-ba* 'not go:PAST-2SG away the school-DIR,' i.e., 'you didn't go (away) to school,' with *nem* in focus position forcing the verbal prefix *el-* out of focus position, and also the discussion of Armenian focus in Comrie (1984).

ON THE SYNTAX OF LITHUANIAN PARTICIPLES

Bill J. Darden
University of Chicago

In Lithuanian the order of major constituents of a clause is determined by principles of pragmatics.[1] This causes problems for theories of syntax based on languages like English, in which word order reflects constituent structure. Not the least of the problems is that there is no surface evidence for a constituent which we could call a verb phrase in Lithuanian.

In this study we will approach Lithuanian using a theory similar to that of autolexical syntax. In this theory syntax, morphology, and semantics are considered to be independent modules with their own rules. Utterances of a language must meet the well-formedness conditions of all modules independently. The rules for morphology discussed by Sadock (forthcoming) have been primarily been concerned with word-internal morphology. To handle highly inflected languages like Lithuanian, Russian, Latin, or Greek, we need a module that treats the aspects of morphology which express the relationships between words in a sentence. Anthony Woodbury (personal communication) has suggested calling this module 'inflection,' which seems appropriate enough. Some things which are normally included under the notion 'inflection' (mood, aspect, plural, etc.) may not always indicate relationships between words. However. they are often synthetically combined with categories with do have this function, or they may play a role in agreement. Aspect and tense at the very least affect the semantic interpretation of whole groups of words. When they are morphologically part of a single system which is reasonably called inflectional, there would be no conflict in including them.

Very little attention has been paid to these aspects of grammar in autolexical syntax. There is a preliminary discussion in Schiller (forthcoming).

In what follows we will attempt to characterize the proper use and interpretation of participles in Lithuanian without the use of any abstract syntactic structures which are not justified in the surface syntax of the language. The semantics will be interpretive, but we would not wish to rule out the possibility of an abstract tier of formal semantics as suggested by Jan Faarlund (forthcoming).

Such concepts as 'subject' and 'object' of a verb are as relevant for Lithuanian as for any other language, and we should give a preliminary definition for them.

Within the inflectional component, the subject of a finite clause is indicated in two ways. It is indicated by the person and number marking of the finite verb, or by a noun phrase in the nominative case. To actually be interpreted as the subject, the NP in the nominative must be pragmatically interpretable as the thing of which the verb is predicated.

From the point of view of the inflectional component, the following two sentences are ambiguous:

(1) Mano mokytojas yra Jonas Petraitis.
My teacher-nom is Jonas Petraitis-nom.
My teacher is Jonas Petraitis
(2) Jonas Petraitis yra mano mokytojas.
Jonas Petraitis is my teacher.

The choice between interpretation as subject or predicate nominative is disambiguated by word order or intonation.

The unmarked indication of the direct object of the verb is the accusative case. The use of case inflection to indicate the object, rather than the use of rigid word order, is not only a different way of indicating this relation—it allows for the expression of additional information about the relationship. The NP complement of the verb can be expressed by other cases. In a theory which treats syntactic trees as primary, and surface case as derived, the oblique objects must be treated as lexical idiosyncrasies of no syntactic significance. When we consider the surface case to be of independent status, however, we find that the oblique case objects are at least as reliable as indicators of the relationships between NPs and verbs as are lexical prepositions.

As a first approximation, we could say that the dative case indicates a negative or positive beneficiary of the action, generally a person affected but not physically moved by the action of the verb.

Genitive objects in Lithuanian indicate that the object is wholly or in part unaffected by the action of the verb. This includes partitive objects, objects of negated verbs, and objects of verbs like *bijoti* 'fear,' *laukti* 'wait for' whose objects might be said to have a semantic role of 'source.'

We can now turn to the syntax of participles.

Inflected participles are used in attributive position and in appositive phrases in Lithuanian (corresponding roughly to the function of restrictive and nonrestrictive relative clauses). We will not discuss these functions except insofar as they are relevant for the discussion of the functions which interest us. Nor will we discuss in detail the use of participles in the reported mood.

The phenomena we will deal with are:

(a.) The choice of a participle in the nominative case vs. a gerund phrase in adverbial phrases. This is well understood and easily accounted for, but it will be useful to discuss in relation to the other problems.

(b.) The choice of 'special participle' or regular participle in the nominative case.

(c.) The choice of gerund or inflected participle in verbal complements.

(d.) Finally, we will discuss in more detail the semantic interpretation of object complements.

We will deal only with active forms of the participles (there is no passive gerund). Examples of the present and past participles (in the nominative, accusative, and genitive singular) and the gerunds are:

		Participle		Special Participle		Gerund
		M	F	M	F	
Pres	Nom	grįžtąs	grįžtanti	grįždamas	grįždama	grįžtant
	Acc	grįžtantį	grįžtančią	X	X	"
	Gen	grįžtančio	grįžtančios	X		X
Past	Nom	grįžęs	grįžusi	=participle		grįžus
	Acc	grįžusį	grįžusią	X	X	"
	Gen	grįžusio	grįžusios	X	X	"

The special participle, which is formed with the suffix *-dam-*, occurs only in the nominative case. In tenses other than the present, there is no special participle, and its function is taken over by regular active participles. As is obvious, the gerund is not inflected for nominal categories, although it is inflected for tense. As we shall see, the gerund is not only less nominal in nature than the participle; it is more verbal. It functions as the verb of nonfinite constructions and can take a subject.

Adverbial Phrases. In adverbial use the choice is between the gerund and the participle in the nominative case. Instead of present active participles, we find the special participle. The participles in the nominative indicate that the understood subject of the action is the same as the nominative subject of the matrix clause. The gerund is used in all other cases. Examples with the special participle are:

(3) Gailiai verkia dukružėlė, nuo močiutės **eidama.** GLJ 313.
Bitterly cries daughter, from mother **go-spec part**
The daughter cries bitterly as she leaves her mother.

(4) Langas atsidarė, **įleisdamas** gaivaus oro. GLJ 313
window opened, **let in-spec part** fresh air.
The window opened, letting in fresh air.

With past participle:

(5) Išėjusi iš miško, Veronika net stabtelėjo. GLJ 313
exit-past part-nom from forest, Veronika even stopped.
After leaving the forest, Veronika even stopped.

As we noted above, the gerund is used when the subject of the secondary action differed from the subject of the main verb. The subject of the gerund may be specified in the dative case (the dative absolute). An example with the dative subject is:

(6) **Žuvus Mindaugui,** susvyravo net jo sukurtos valstybės pamatai. (LI 59)

die-past-ger M-dat, was shaken even by him founded state-gen foundations
When M perished, the foundations of the state he founded were shaken.

When the subject is not specified, it is interpreted pragmatically. The possibility of the subject of a gerund being absent and interpreted from context corresponds roughly to the possibility of having no pronoun in a finite subordinate clause—i.e., we could interpret it as the same phenomenon as pro-drop. The understood subject is commonly in the main clause (in any case but the nominative). An understood subject coreferential with the accusative in the main clause is:

(7) Mus ir apšvito **bešnekučiuojant.** (L-AKŽ)
Us-acc and shined-on **talking-ger**
Day broke while we were talking.

Coreferential with the main clause dative:

(8) Tau yra geriau su viena akimi įeiti į gyvenimą, nekaip **abi akis turint,** būti įmestam į ugnies pragarą. (Math 18: 9) B 1961.
You-dat is better with one eye to enter heaven, than **both eyes have-ger**, be cast into hell.
It is better for you to enter Heaven with one eye than, having both eyes, be cast into Hell.

(9) **Sunkiai dirbant,** arkliams reikia grūdo. LKG II 390
Hard work-ger, horses-dat need grain-gen.
When they work hard, horses need grain.

The gerund from an impersonal verb, of course, is systematically excluded from having a subject, as in:

(10) Grybai tik **lyjant** ir auga. GLJ 320
Mushrooms only **rain-ger** (emphatic particle) grow.
Mushrooms grow only when it rains.

If the correspondence between the subject of the participle and the main clause subject is only partial, the gerund is used:

(11) Vakar, **kertant rugius**, buvau labai sušilęs. LKG II 390
Yesterday, **cut-ger rye,** (I) was very heated.
Yesterday, when (we) were cutting rye, I got very hot.

In the above sentence, the gerund phrase would be better translated "during the rye harvest." It does not specifically indicate that the subject is "we," only that the subject does not precisely coincide with the subject of the matrix verb. We pick up the fact that the speaker engaged in the harvest from the implication that there is some connection between the cutting of the rye and his getting hot.

In these constructions, the past participles and gerunds indicate that the action of the adverbial phrase precedes the action of the main verb. The present participles and gerunds generally indicate an action which is simultaneous with the action of the main verb. For other temporal relationships, either the participles or the dative absolutes can follow a preposition, e.g.:

(12) **Prieš gaidžiui pragystant,** tu tris kartus manęs užsiginsi. (Mat.. 26:75) B 1961.
Before cock-dat crow-ger, you three times me will deny.
Before the cock crows, you will deny me thrice.

(13) Nėra žinių, kurių sričių kunigaikščiu buvo Mindaugus, **prieš paimdamas į** savo valdžią visą Lietuvą. (LI 49)
is-not knowledge, of what area prince-ins was M-nom, **before take-spec part** in his rule all Lithuania-acc
There is no indication as to what area Mindaugas was prince of, before he took all of Lithuania into his realm.

Such constructions are rarer with *iki* or *lig* 'until, while.'[2]

(14) Vėl dėdė nieko nebedirbo, tik ištisas valandas stovėjo, stovėjo ties langu, **iki kojoms nutirpstant.** LKG III 534

more Uncle nothing not do, only whole hours stood, stood before the window **until legs-dat go numb-gerund.**
Uncle didn't work anymore, only stood, stood for whole hours in front of the window until his legs would go numb.

(15) **Ligi nesutemus**, reikėjo vaikščioti. LKG III 534.
While not get dark-gerund, be necessary to take a walk.
While it hasn't gotten dark, one must take a walk.

(16) Susikalbėkim, berneli mano, **lig jauni bebūdami.** LKG III 534
Lets talk, boys my, **while young be-spec part**.
Lets talk, my boys, while we are still young.

Special vs. Regular Participles. We will now turn to the discussion of the choice between special participles and regular participles in other environments. The special participles occur in a variety of headless relative constructions which function as adverbs of time, manner, or place, or as indefinite arguments of the verb. Examples are:

(17) Bet aš niekaip negaliu užtikti pinigų, nors ieškau **kur nutrūkdamas** ir seku Saliamutę **kaip tik galėdamas.** LKG II 357
But I nohow can't find money, although I search **where(ever) take a break-spec. part** and follow S **as much as can-spec. part**.
But I can't find the money no matter what, although I search wherever I take a break, and follow S. as much as I can.

(18) Dirbu **kaip galėdamas** GLJ 314
I work **how can-spec part**
I work as well as I can.

(19) Davė knygas **kam norėdamas.** (informant)
He gave books **to whom want-spec part**
He gave the books to whomever he wanted to.

(20) Sakyk, **ką norėdamas.** (informant)
Say what(ever) you want to.

Compare:

(21) Sakyk ko nori.
Tell what-gen want-2nd sing, indicative present
Tell (me) what you want.

(22) Sakyk, ką **išmanydamas.** LKG II, 357
Say what-acc **understand-spec part**

This sentence, not glossed in the grammar, is problematical. Since these constructions are apparently not used as embedded indirect questions, (cf. 20, 21, 30), it should mean something like "Tell whatever you can figure out," or " Say what comes to mind." (This was confirmed by a native speaker.)

(23) Kuo **turėdama**, mane valgydino (DLKŽ 913)
What-ins (she) **had-spec part** me (she) fed
She fed me what(ever) she had.

This sentence poses still another problem. The instrumental *kuo* has its case governed by the main verb *valgydino,* rather than by the participle *turėdama.* This provides an interesting intersection of the semantic, inflectional, and syntactic components. In their inflection, both *kuo* and *turėdama* are independent parts of the main clause. However, their word order is fixed, conforming to the fact that they constitute a clause at the semantic level. This phenomenon (the case of the *k*- pronoun being determined by the main clause) is not mentioned in the standard grammars.

In all the other examples, the form of *k*- pronoun is ambiguous—it would be the same in both the upper and lower clause. In sentences (19) and (20) the *kam* and *ką* seem to be inflectionally part of the main clause, but they could be the objects of understood infinitives *duoti* 'give' and *sakyti* 'say.' An attempt to construct an unambiguous example (24) resulted in a sentence judged 'so-so' by a native speaker.

(24) ?Davė knygas **kam žinodamas?**
He gave books **to whom know-spec part**

This is an interesting aspect of these constructions, particularly from the point of view of autolexical syntax, but the point of this paper is that all the uses of the special participle have the semantic structure of headless relatives (relatives with no specific NP in the main clause which they define). This is also true of the adverbial uses. Adverbs of time or place can be interpreted as having the structure of headless relatives. In Old Slavic, Sanskrit, and Greek, the relevant adverbs are relative formations. "When" or "where" can be interpreted as "at the time when," "at the place where."

The regular participles in the present can represent almost any other reduced embedded clause with the subject coreferential with the main subject and the clause in the present tense.

An example which is superficially similar to the adverbial special participles is:

(25) Ir štai jiems pasirodė Mozė ir Elijas, **besikalbą su juo.** Mat. 17:3. B 1961.
And suddenly to them appeared Moses and Elijah, **talking-pres. part with him.**

Here the participial phrase is part of the subject, rather than being an external phrase whose understood subject is identical with the main subject. In another translation (B 1972) the participial phrase is replaced by a relative clause: *kurie kalbėjosi su juo,* 'who were talking with him.'

Regular active participles are very common with verbs of cognition and communication. Examples are:

(26) Mykoliukas nenusimanė Severjai **griežiąs** LKG II 358
M-nom not understand-past S-dat **play-part-nom**
M didn't understand that he was playing for S.

(27) Jis sakosi **žinąs.** Schmalstieg 88
He say-reflexive **know-part-nom**
He says that he knows

(28) jis dėjosi mano kalbos **negirdįs** LKG II 359
he pretended my speech **not-hear-part-nom**
He pretended that he didn't hear me.

(29) Išrinktas didžiuoju Lietuvos kunigaikščiu, Švitrigaila tuojau pasirodė **eisiąs** Vytauto pėdomis.
Selected-ppp-nom great-ins L-gen prince-ins, Š-nom right away showed **go-fut-part-nom** V-gen footsteps-ins.
Once he was selected Grand Prince of Lithuania, Švitrigaila immediately showed that he would follow in the footsteps of Vytautas.

An example in an embedded question is:

(30). jie nežino, **ką darą.** (Luke 23: 34) B 1961, 1972.
they not know, **what-acc do-part-nom.**
They know not what they do.

Participles are somewhat rarer with verbs of perception:

(31) .. pamatė **apgauta esanti** LKG II 351
(she) saw **deceived-nom be-part-nom**
she saw that she had been decieved

(32) ..tu jautiesi **esąs nuskriaustas** LKG II 351
you feel **be-part-nom cheated-nom**
you feel that you are cheated.

The present active participle can also be used with *likti* 'remain, keep.'

(33) Senutė liko **beverkianti.**
The old woman remained (behind) **crying.**

There are several uses which call specifically for the past active participle:

(34) Jau liovėsi sniegas **kritęs** (dial) LKG III 177
Already stopped snow-nom fall-past-active-nom.
The snow has already stopped **falling (past-active)**

(35) Artojas metė **aręs.** LKG III 177
The plowman stopped **plowing (past active)**

(36) Girdėdama vyrų kalbą, neiškentė **neatsiliepusi.** LKG III 178
hear-spec part men's speech, she could not stand **not-answer back-past-active.**
Hearing the men's speech, she could not resist answering back.

(37) dirbk **neplepėjęs!** LKG III 178.
work-imp **not-waste time past-active.**
Work without wasting time.

There is at least one type of construction with an apparent headless relative which takes the regular participle:

(38) Nėra **kas duodąs.** GLJ 314
not-is **who give-pres. part.**
There is no one who will give.

There are two possible readings of the surface structure. In one interpretation *kas duodąs* is the subject clause. In that case *duodąs* cannot have an understood subject identical to the main subject, since it is part of that subject. In another reading, the clause is promoted to the main clause, with *kas* as subject and *duodąs* in predicate nominative position. Special participles do not occur in such positions.

Sentence (39) causes special problems:

(39) Nėra kas darą. GLJ 314
not-is what-nom do-part-nom neuter
There is nothing to do.

This sentence is syntactically well formed for two reasons. The pronoun *kas,* when it means 'what,' can take "neuter" agreement (lack of agreement). (We might more precisely say that the pronoun *kas* can be interpreted as the predicatee of a neuter adjective or participle as long as it can be simultaneously interpreted as meaning 'what'). Furthermore, predicate adjectives have neuter forms with infinitives, when the semantic object of the infinitive shows up in the nominative case (cf. the discussion of (54)).

(40) Sita knyga sunku skaityti.
This book-nom hard-neuter read-infinitive
This book is hard to read.

The fact that *kas* is nominative, and the fact that the participle is neuter, can be related to these facts. Timberlake (1974) suggests that this is a direct realization of an object in the nominative case in an impersonal construction. Ivan Sag (personal communication) has suggested that the syntax of Lithuanian could be described as allowing neuter forms in the predicate adjective when there is no noun in the nominative to agree with, or when there is a noun, but there is a gap in the object position of an active infinitive or participle in the construction.

We can see that it would be very difficult to place limits on where an active participle can occur representing a semantic embedded clause in the present or past *realis*, with subject the same as the subject of the main clause. The only thing that blocks this interpretation is an overt complementizer, which sets up a clause boundary, as in:

(41) Visi pasakoja, **kad tu esanti** pirmoji šiame miške krembliautoja... LKG II 364
All say, that **you be-pres part** first in this forest mushroom gatherer (female).

Everyone says that you are the premier mushroom gatherer in this forest.

Here the clause introduced by *kad* is participial because it is in the reported mood. The participle predicates the subject of its own clause.

Complement Clauses. We can now turn to the discussion of the choice of gerund or inflected participle in verbal complements where the subject of the complement differs from that of the main clause. In these phrases, the subject of the participle or gerund is in the case expected of an object of the matrix verb. The inflected participle agrees with this NP in case. It is useful to separate complements into three semantic classes. (1) Propositional complements, complements of verbs of communication and cognition, which represent single clauses in the semantics. (2) Event complements, complements of verbs of perception and verbs like *laukti* 'wait for,' which may take single-clause or NP plus participial phrase. (3) Complements of verbs like 'find,' 'leave,' which clearly take NP plus participial phrase.

In the modern language, propositional complements, when not independent clauses in the surface, are expressed with gerund phrases. Complements of the third type, on the other hand, rarely are expressed with gerunds. They are typically an NP plus an inflected participle. The event predicates, being ambiguous, show both structures. This means that a gerund phrase typically represents an S (a single clause) in the semantic structure. This is also true of the adverbial uses of the gerund, which represent whole clauses in the semantics. The NP plus inflected participle, on the other hand, typically represents two elements in the semantic structure: a noun phrase object and a reduced clause. The reduced clause is what in GPSG would be an S/NP, an S which has lost an NP. To identify the missing NP, we look for the NP with which the participle agrees in case—the object of the verb. Case thus serves to link the participle to an NP which has some other function in the sentence. The interpretation of a participial phrase as NP/S parallels the interpretation of attributive and appositive participial phrases.

Examples of propositional complements are:

(42) Traidenis sakė **jį esant pastatytą** jo paties širdy LI 61
Traidenis said **it-acc be-ger put-ppp-acc** his very heart-loc
Traidenis said it was stuck in his very heart.

(43) Kai kurie mokslininkai spėja Mindaugo **sostinę buvus Liškiavą ar Kernavę.** LI 77
Certain scholars believe M-gen **capital-acc be-past-ger L-acc or K-acc**
Certain scholars believe that Mindaugas' capital was Liškiavas or Kernavas.

(44) aš pažinau **tave esant kietą žmogų** (Math 25: 24) B-1961
I knew **you-acc be-ger hard man-acc**
I knew you to be a hard man.

(45) Kaip jauti **tėvą pasielgsiant?** GLJ 321
How do you feel **father-acc behave-act-fut ger**
How do you feel that father will react?

The use of gerunds in these complements represents a historical innovation. The archaic language preserved the Indo-European pattern of inflected participles with agreement in case:

(46) žinau **tave gražią moteriškę esančią** Bible (1579-90) GLK II 363
I know **you-acc beautiful woman-acc be-part-acc**
I know you to be a beautiful woman.

(47) Kuomi sako **mane esantį** žmonės. Mark 8:27 (1816) GLK II 363
What-ins say **me-acc be-part-acc** people-nom
Who do the people say that I am?

A modern version of the above passage is:

(48) Ką žmonės sako **mane esant.** (B 1961)
What-acc people say **me-acc be-ger**

Examples of event complements, which may be either gerund phrases or NP plus participle are:

(49) Išgirdom graudžiai **verkiant jaunąją mergužę**. GLJ 321
We heard bitterly **cry-ger young girl-acc**
We heard a young girl crying bitterly.

(50) .. išgirdo **bitutę atgaudžiančią** LKG II 362
heard bee-acc **buzzing-part-acc**
(he) heard a bee buzzing toward him.

(51) Moteriškė... pamatė **vyrą ateinant**. GLJ 321
Woman saw **husband-acc come-ger**
The woman saw her husband coming.

(52) Mačiau **brolytį grįžtantį** LKG II 362
I saw **brother-acc return-part-acc**
I saw my (little) brother returning.

(53) Laukiu sūnaus **grįžtant/grįžtančio** DKS 210
I wait-for son-gen **return gerund/part-gen.**
I am waiting for my son to return.

(54) Kęstučio gatvėje girdėti **grojant orkestras.** LKG II 395
K-gen street-loc hear-inf **play-ger orchestra-nom**
On K street one can hear an orchestra playing.

This sentence requires special comment, because the subject of the gerund phrase is in the nominative case. In Lithuanian the semantic object of an infinitive in an impersonal construction can be realized in the nominative case. Thus it would be possible to say *Gatvėje girdėti orkestras.* "On the street one can hear an orchestra," or "... an orchestra can be heard." There is a long history of discussion of these formations, centering on whether the NP in the nominative is to be considered the subject. A recent article with considerable bibliography is Ambrazas (1987). A much more extensive discussion (and an opposite conclusion) is found in Timberlake (1974). For our purposes the important fact is that the matrix sentence allows nominative case in this structure. The phrase *grojant orkestras* can be considered a gerund phrase interpreted semantically as a clause which is an argument of *girdėti*. Its subject gets its case from the matrix sentence—in this instance the nominative case.

This sentence forms a minimal triple with (3) and (25). In (54) we find a gerund in an NP which is an S as the (probable) subject. In (25) we find a regular participle as an attribute inside an NP subject. In (3) we have an adverbial expression external to the subject NP, expressed with the special participle.

We can now turn to predicates of the third class. The Academy Grammar (LKG II:393) says that the gerund is rarer with the verbs *sučiupti, sugriebti, pagauti, nutverti* 'catch,' *sutikti* 'meet,' *užeiti* 'find,' *palikti* 'leave.' Regular examples of inflected participles with verbs of this type are:

(55) pavijome **ją einančią** vieškeliu. Vienuolis 7
we caught-up-with **her-acc go-part-acc** road-ins
We overtook her walking along the road.

(56) Radau brolelius **begulinčius** LKG II 362
I found brothers-acc **lie-part-acc-pl**
I found my brothers lying down.

An apparent example of a gerund complement with *rasti* 'find' is (57), contrasting with (56):

(57) Parėjo namo ir rado **sūnų užgimus**. GLJ 321.
Came home and found **son-acc born-past-ger**
Prišel domoj i uznal, čto rodilsja syn.
He came home and found that a son had been born.

In fact, gerunds are often found with *rasti*. However, the Russian translation of (57), provided in the grammar, indicates that the verb should be translated as 'find out' rather than 'find' in the sense of

'come upon.' This makes the complement a propositional complement, for which the gerund is normal.
Nevertheless, we do find:

(58) Aš jį palikau **begulint/ begulintį** DKS 210
I him-acc left **lie-gerund/ part-acc**
I left him lying down.

(59) sučiupau šešką, **kiaušinius beėdant.** DLKŽ 752
I caught polecat, **eggs-acc eating-ger**
I caught a polecat (while he was) **eating eggs.**

One solution would be to admit the existence of an innovative style which allows gerunds in any verbal complement.

We know that the gerund originated in dative absolutes: it is the phonologically reduced form of an old dative singular of the participle. At the oldest level, all participial complements were inflected participles which agreed with NPs in case. With the loss of the ending, the gerund became an uninflected form which occurred with an oblique case subject. Semantically, it represented an adverbial clause. There are two degrees of spread into other constructions. It has spread mainly to structures which are most similar in form and meaning to the dative absolute—participial complements with oblique case subjects which represent single clauses at the semantic level. In less conservative styles, it may spread to formations which represent true direct objects plus participial phrases. This could be presented in the following table:

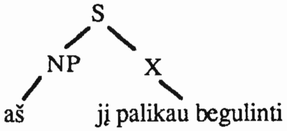
	Propositional complements		**Event Complements**		**NP + S complements**	
Archaic	NP-case	part-case	NP-case	part-case	NP-case	part-case
Normal	NP-case	gerund	NP-case	part/gerund	NP-case	part-case
Max innovative		NP-case gerund for all three constructions				

An alternative approach to the analysis of complements of the third class, which will be explained below, would be to treat the structures with gerunds as a different type of structure—the gerund representing an independent clause with its understood subject coreferential with the object of the verb.

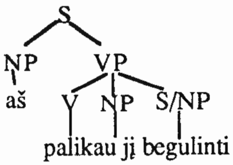
Semantic Interpretation. We have so far been discussing verbal complements as if it were obvious what they are. In the absence of evidence for constituent structure, this may not be justified. It is obvious that the subject in the nominative case plays a special role in the sentence structure of Lithuanian. There seem to be no constraints on where within a clause a participle can occur and agree in case with the subject. This is analogous to the fact that in a syntactic analysis of a language like English, the subject commands everything within its clause.

Agreement in case with other arguments of the verb is more restricted. For a participle to agree in case with the object, for instance, it must have a close semantic relationship with the object. Both must be part of the same noun phrase or within the scope of the verb.

We can use these facts to build a set of hierarchical inflectional relations which are not unlike a syntactic tree. We will start with sentence (58) in the option with participial case agreement: *Aš jį palikau begulintį*. We first set off the nominative:



The case agreement between *jį* and the participle justifies linking them in a single constituent, and the fact that the pronoun is in the accusative case marks it as the direct object of the verb. We then have the inflectional equivalent of a VP:



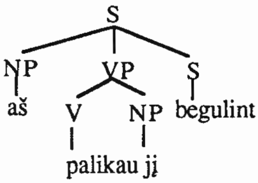
This is not of course a semantic interpretation. Subject and object are not semantic entities. We are used to thinking of subject and object as syntactic entities, but if syntax is defined as sentence structure relating to surface word order, then subject and object are not syntactic items in Lithuanian. They are, however, very clearly items of sentence structure. What we want to say is that the relationship between a verb and, say, its direct object, is just that—a relationship. This relationship is realized through word order in English, and so can be abstracted from that order. In languages like Lithuanian, it is realized through inflection, and can be abstracted from that. Insofar as the relationship is structural, rather than purely semantic, and insofar as the relationship is roughly the same in the two languages, then it cannot be syntactic in the narrow sense—in either language. What we need is a level of analysis which is sentence structure abstracted from surface order and inflectional morphology. When a linguist analyzes a language like Lithuanian with a 'syntactic' tree which reflects and abstract arrangement of constituents (cf. George Fowler's 1987 treatment of Russian case), he is creating this level of sentence structure and symbolizing the hierarchical relationships in a tree structure. The mistake that one easily makes is identifying this tree with an analysis of the order of constituents.

Jan Faarlund does have such a level of analysis, where he treats the relationship between a verb and its clausal arguments in the manner of relational grammar. He, however, calls this a tier in the semantic representation. We could suggest that he keep the level, but change the label.

One might claim that in admitting the existence of a level of sentence structure which has hierarchical relationships not expressed in word order, we are essentially doing slightly abstract syntax—just giving it another label. In this case, however, we are not using this abstract level to govern the use of case. We are using it to embody the information conveyed by inflectional structure. It is this structure which is most easily related to a more purely semantic level. It will be interesting to see how far we can go without referring to structures not justified by the inflection.

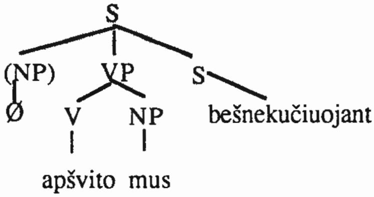
Obviously, the identity of the subject and object of a verb is essential for the interpretation of the semantics. What a verb predicates of its subject and objects is an essential part of its lexical semantics. This structure, combined with the lexical semantics of the individual items, will yield an interpretation of *Aš jį palikau.* The participle can be interpreted as representing a clause at the semantic level. Because of its case, we assume that the subject of that clause is coreferential with the direct object and the clause is in the scope of the verb.

What then of *Aš jį palikau begulint,* the option with the gerund? For this structure we can assume that a single-clause predicate complement is ruled out on semantic grounds. We have no inflectional evidence that the gerund has a particular subject, or that is within a VP. If we limit ourselves to the structure that is justified in the surface, we have:



If we leave the structure like this, we are saying that the gerund in this sentence is adverbial and does not differ in function from the gerund in (7):

(7) Mus ir apšvito **bešnekučiuojant.** L-AKZ
Us-acc and shined-on **talking-ger**
Day broke while we were talking.



This is at the very least an unusual position to take. A theory which takes syntactic trees as primary would almost certainly treat (58) in both options (with gerund or participle) as having a single structure, with the gerund/participle in the verb phrase. The choice between gerund and participle would be handled by making agreement in case optional. The gerund would be a participle with no agreement. The gerund in (7) would be treated as adverbial.

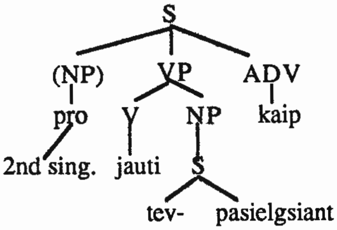
There is apparently nothing in the inflectional endings, in the word order, or in the intonation to differentiate the use of the gerund in (58) from that in (7). If there is a difference between adverbial gerunds and gerunds in verbal complements, it is neutralized in this structure. The position after the verb and the lack of an intonational break indicates that the gerund is an integral part of the rheme in both cases. That may be enough.

It turns out that we can adequately interpret *Aš jį palikau begulint* without assuming anything about the position of the gerund in the verb phrase. The gerund represents an S. Since its subject is not specified in the sentence, it must be redundant in the context. It cannot be coreferential with the main clause subject, so in this sentence it must be coreferential with the object. We thus know that it was *jis* (the object) that was lying down, and that his state of lying down obtained at the time that I left. We do not need to know any more.

Verbs which take only propositional complements have similar surface structure, but cannot be interpreted in the same manner. An example is (45):

(45) Kaip jauti **tėvą pasielgsiant?** GLJ 321
How do you feel **father-acc act-fut ger**
How do you feel that father will react?

Here the semantics of the verb eliminates the possibility of interpreting *tevą* as the semantic direct object and the gerund as an independent clause. *tėvą pasielgsiant* is then interpreted as a nominalized S. The fact that *tėvą* is accusative indicates that the S is an object clause. This justifies a shallow semantic representation:



The interrogative adverb is left dangling from the main S node here because we have no inflectional indication as to where it goes. Pragmatically, it would be put in the gerund phrase.

When the gerund phrase has a predicate nominative or adjective, as in (44), there is an inflectional indication that the gerund is part of the object clause (and not an independent gerund phrase). The predicate nominal or adjective is in the case of the object:

(44) aš pažinau **tave esant kietą žmogų** (Math 25: 24) B-1961
I knew **you-acc be-ger hard man-acc**
I knew you to be a hard man.

For verbs of perception, we have two alternate surface structures, exemplified by:

(51) Moteriškė... pamatė **vyrą ateinant**. GLJ 321
Woman saw **husband-acc come-ger.**
The woman saw her husband coming.

(52) Mačiau **brolytį grįžtantį** LKG II 362
I saw **brother-acc return-part-acc.**
I saw my brother returning.

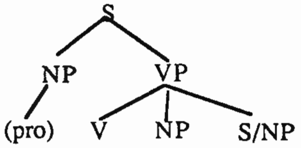
Assuming that what we said earlier was right, and we can interpret *vyrą ateinant* in (48) as a nominalized clause, then (51) and (52) represent two equally legitimate ways of packaging the complements of verbs of perception. These verbs obviously allow single NPs as objects, as in: *mačiau brolytį* 'I saw my brother.' When a scene is perceived in which an entity engages in an activity, it is possible to represent this with the entity as direct object with an additional complement specifying the activity, or with a clausal complement with the entity as subject. In Russian the normal complement is clausal, introduced by the complementizer *kak.*:

(60) Ona videla kak muž prixodil.
She saw comp. husband-nom arrive
She saw her husband arriving.

This corresponds to our interpretation of (51), except that (51) has a nominalization which has no independent clausal status.

Because the verb allows a direct object NP, (51) is actually ambiguous as to its surface structure. It could be interpreted as having an NP object and an independent gerund. Again, from context, the subject of the gerund would be understood in this case as the same as the object NP. It would mean something like "She saw her husband as he came." Again, it makes little difference what interpretation we take.

Sentence (52) however has an interpretation similar to that of *Aš jį palikau begulintį*, analyzed above.



We have so far not attempted to decide the structural relationship between the participle in the accusative and the direct object of the main verb. The participle could be an independent complement of the verb, or it could be interpreted like a nonrestricted relative clause. This ambiguity does not have to be resolved. There is little relevant difference between "I saw my brother returning" and "I saw my brother, who was returning." There is more difference in English than in Lithuanian. This becomes particularly obvious if we take a sentence like (25) and convert it into a sentence with an active verb of perception.

(25) Ir štai jiems pasirodė Mozė ir Elijas, besikalbą su juo. Mat. 17:3. B 1961.
And suddenly to them appeared M & E, talking-pres. part with him

becomes:

(61) Jie pamatė Mozę ir Eliją, besikalbančius su Jėzumi
They saw M-acc and E-acc, talking-acc with Jesus-ins

As was the case with (25), the participle can be replaced by relative clause without a change in meaning:

(62) Jie pamatė Mozę ir Eliją, kurie kalbėjo su Jėzumi.
They saw M-acc and E-acc, who were talking with Jesus-ins.

The semantic interpretation of nonrestrictive relatives as totally independent statements (cf. McCawley 1988:426) is not justified in this case. The relative clause is clearly within the scope of the verb, i.e., it is part of the scene which is perceived.

Conclusion. What we see then is that it is possible to describe and interpret the structures of participial and gerund phrases without postulating an abstract level of syntax which is not justified by surface relationships. In particular, we find that when information about constituent structure is lacking in the surface, that structure may be irrelevant for the semantic interpretation. Thus we do not have to decide whether a gerund is part of a constituent that corresponds to a verb phrase, or is an independent adverbial; we do not have to decide whether an inflected participle is an attributive, appositive phrase, or is some other kind of verbal complement.

In these cases, lack of specific information at the inflectional level may result in vagueness at the semantic level, rather than ambiguity. We cannot say, however, that there is no ambiguity at the inflectional level. Obviously there is an ambiguity between the two uses of *rasti* 'find' and 'find out'. So *Jis rado žmoną gulint* could mean 'He found his wife (when she was) lying down,' or 'He found out that his wife was lying down.' There is similar ambiguity between the senses of *matyti* 'see' or 'comprehend' and *girdėti* 'hear (an occurrence)' or 'hear (news)'. In the second sense, *matyti* and *girdėti* take propositional complements. Since we have argued that the absent subject of a gerund is picked up from context, as would the comprehension of the referent of a pronoun, there can obviously be cases of ambiguity of reference. Nevertheless, the lack of need for a syntactic structure that is not justified in the surface is striking.

We have of course used fairly simple sentences. In more complex sentences there would at least be adverbial expressions in the accusative or nominative case. This would require much more use of pragmatics to differentiate between more than one use of the same surface case. There is nothing in principle that would make this difficult. After all, Lithuanians have managed it for centuries.

SOURCES

B 1961 = Naujasis Testamentas. London: The British and Foreign Bible Society.
B 1972 = Šventasis Raštas: Naujasis Testamentas. Boston: Amerikos Lietuvių R. K. Kunigų Vienybę.
DKS = Dambriunas, Klimas, and Schmalstieg 1966.
GLJ = Ambrazas, et al. 1985.
L-AKŽ = Piesarskas and Svecevičius 1979
LI = Šapoka, A. 1936. Lietuvos istorija. Kaunas: Knygų Leidimo Komisijos leidinys.
LKG = Ulvydas et al. 1971-6
Schmalstieg = Schmalstieg 1987.
Vienuolis = Vienuolis, A. 1982. Iš mano atsiminimų. Kaunas: Šviesa.

REFERENCES

Ambrazas, V. 1979. Lietuvių kalbos dalyvių istorinė sintaksė. Vilnius: Mokslas.
_____. 1987. Indogermanische Grundlage des Dativus und Nominativus cum Infinitivo im Baltischen. IF 92:203-219.
_____. et al., eds. 1985. Grammatika litovskogo jazyka. Vilnius: Mokslas.
Dambriunas, L., A. Klimas, and W. Schmalstieg. 1966. Introduction to Modern Lithuanian. Brooklyn: Franciscan Fathers.
Faarlund, Jan. Forthcoming. "Autostructural Analysis," paper presented at the Workshop on Autolexical Syntax, University of Chicago, 1989. The proceedings have been submitted to Mouton for possible publication.
Fowler, George. 1987. The Syntax of the Genitive Case in Russian. University of Chicago Dissertation.
Kruopas, J., et al., eds. 1972. Dabartinės lietuvių kalbos žodynas. Vilnius: Mintis.
Lyberis, A. 1962. Lietuvių-rusų kalbų žodynas. Vilnius: Valstybinė Politinės ir Mokslinės Literaturos Leidykla.
McCawley, J. D. 1988. *The Syntactic Phenomena of English.* Chicago: University of Chicago Press.
Piesarskas, B. and B. Svecevičius. 1979. Lietuvių-anglų kalbų žodynas. Vilnius: Mokslas.
Sadock, Jerrold M. Forthcoming. *Autolexical Syntax: A Theory of Parallel Grammatical Representations.* University of Chicago Press.

Schiller, Eric. Forthcoming. "The Case for Autolexical Case," paper presented at the Workshop on Autolexical Syntax, University of Chicago, 1989. The proceedings have been submitted to Mouton for possible publication.
Schmalstieg, William R. 1987. A Lithuanian Historical Syntax. Columbus, Ohio: Slavica.
Timberlake, Alan. 1974. The Nominative Object in Slavic, Baltic, and West Finnic. Munich: Otto Sagner.
Ulvydas, K. et al., eds. 1971-76. Lietuvių kalbos grammatika, Vols. II, III. Vilnius: Mintis.

NOTES

[1] There is a discussion of the principles of order of major constituents in GLJ: 671-680. For our purposes it is enough to note that all permutaions of S V O are permissible.

[2] I have found one case (from a folk tale and therefore not reliably part of the standard language) where the selection of the participle rather than the gerund seems to violate the rules given above:

(i) Jo buvo įsakyta nieko nedaryti nei motinai, nei kūdikiui, **ligi sugrįšiąs namo.**
LKG II 357
He-gen was ordered-neuter nothing do-infinitive not mother-dat, not child-dat, **until return-fut-part-nom home.**
He ordered (someone) to do nothing to mother or child until he should return home.

This is an impersonal passive with an agent in the genitive which must be regarded as coreferential with the understood subject of the participle. The participle is in the masculine nominative singular, but there is no surface nominative with which it can agree. Here it is possibly crucial that the participle is in the complement of a verb of communication. Participles in the nominative case are commonly used to replace finite verbs in indirect speech. This usage is called 'indirect mood' in modern grammars. Thus we can have:

(ii) Sako, (kad) Jonas atėjęs.
he says that Jonas come pap
He says that Jonas has come.

(iii) Sako, kad (jis) atėjęs
He says that he has come.

The third person pronoun would not normally be used when the reference is to the only other argument of the verb, the subject. If we interpret *ligi sugrįšiąs namo* in the above sentence as an instance of indirect mood with a zero pronominal subject, then it becomes grammatically understandable.

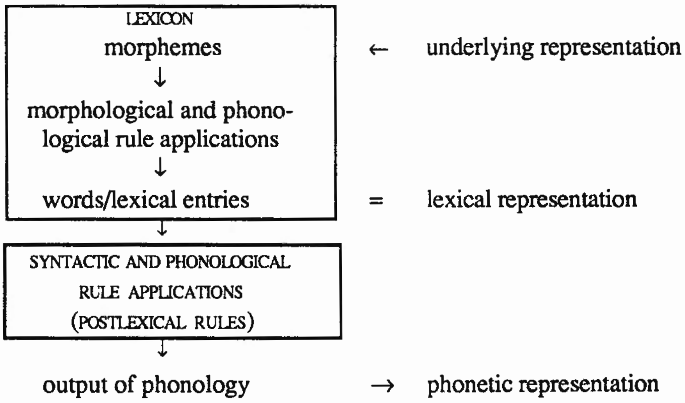
TOWARDS A LEXICAL PHONOLOGY OF CHUVASH: INFLECTION AS DERIVATION

Michael Dobrovolsky
University of Calgary

Introduction. This paper continues an analysis of Chuvash within the framwork of Lexical Phonology (LP) begun in Dobrovolsky 1989. Chuvash is a highly divergent Turkic language spoken in the Chuvash ASSR, some 500 miles east of Moscow and spoken by some 1.3 million people, of which perhaps 475,000 are monolinguals (USSR short statistics, 1984).

Example 2 presents a schematic representation of LP following Mohanen 1986. In the theory, words are constructed in a component of the grammar called the Lexicon, where word-formation rules and relevant phonological rules interact at appropriately specified levels or strata to produce so-called lexical representations — fully formed words after all morphophonemic processes have applied — that are the input to the syntactic component of the grammar. Postlexical rules operate on syntactic representations to deal with phonological variation that is essentially allophonic in nature, that may apply across word boundaries, and that is claimed to be non-neutralizing and exceptionless.

(1) Schematic Representation of Lexical Phonology


LEXICON
morphemes
← underlying representation
morphological and phonological rule applications
words/lexical entries
= lexical representation
SYNTACTIC AND PHONOLOGICAL RULE APPLICATIONS (POSTLEXICAL RULES)
output of phonology
→ phonetic representation

In this paper I will limit myself to a specific issue that arises in Chuvash LP when derivation — specifically compounding — appears to encounter the inflectional component of the grammar.

1. Levels. Both Lexicalist (Scalise 1986) and LP models of the lexicon assume that there may be principled distinctions among the rules of derivation, compounding, and inflection. To date, LP has been less rigid about this than Scalise'gs hypothesis, which claims that these three components are an obligatory part of the morphology. Scalise (1986: 101) allows that the relations among the three word-formation components (derivation, compounding, inflection) may be determined on a language-specific basis. Others, notably Anderson 1982, propose that inflection takes place in the syntactic component of the grammar, since inflection is that which is relevant to the syntactic relations among words. The former model provides derived forms with inflectional endings and then selects the appropriately inflected word as the syntactic structure requires. The latter model assumes that case endings are phonological trappings that are determined by primary syntactic relations (i.e., that 'case' is not a primitive).

First, let us establish that there is a need to distinguish an inflectional from a derivational level in Chuvash. This indeed seems to be the case, as derivational forms do not show evidence of any rules of affrication or consonant deletion like those found with inflected forms (such as, for example, a rule that affricates a case affix-initial *t* to *č* after affixes ending in *n*, as in (2), where the assimilation of the nasal to the point of articulation of the suffix inital C also occurs.[1] It should be noted in 2(c), though, that the imperative inflection does not trigger t-affrication.

(2)	(a)	a derivational form (no Affrication)
	pɨar	'hide' (transitive)
	pɨtan	'hide o.s.' < pɨta+n
	pɨtantar	'cause to hide o.s.'
	(b)	Inflection (Affrication)
	av	'home'

avsem	‘home+Pl.’
avsenče	‘home+Pl.+Loc.’ < av + sem + te < av + sem + re[2]
(c)	Imperative inflection
vakla	‘make small’ < vak ‘small’
vaklan	‘chop up’
vaklantă-r	‘let it be chopped up’

A second example: there is a rule of *n*-deletion that is found at the inflectional level. This rule deletes a suffix-initial *n* when it is preceeded by a consonant-final stem, as shown in (3).

(3)	laša	‘horse’	ĕś	‘affair’
	lašanăn	‘horse+Gen.’	ĕśĕn	‘affair+Gen.’ < ĕś + nĕn
	lašana	‘horse+Dat./Acc.’	ĕśe	‘affair+Dat./Acc.’

In derived forms, suffix-initial *n*s are not deleted postconsonantally.

(4)	vula	‘read’	→	vulană	‘(the) reading’
	piś-	‘cook’	→	piśnĕ	‘ripe, mature’
	pĕl	‘(to) know’	→	pĕlnĕ	‘(I was) a knower’

Finally, forms inflected for possession are also characterized by the following rule of vowel-deletion: a suffix-initial vowel is preserved when it is full; when the initial vowel of the suffix underlying form is reduced, it is deleted and the stem vowel is preserved. The rule is thus: conserve a full vowel nucleus; when there are two full vowels in succession conserve the rightmost one.

(5)	ulma	‘apple’	→	ulmi	‘apple+2PPoss’ (UF: ulma+i)
	ulma		→	ulmam	‘apple+1PPoss’ (UF: ulma+ăm)

Verb stems, however, are derived by the affixation of various mood affixes to the base. I am assuming this is derivation since these stems can ultimately be substantivized by the affix *-mA*. If the bases are not substantivized, they are then subject to inflection for person and number. What is interesting about this verb-stem derivation is that there is a rule of vowel deletion that in some instances conserves the affix-initial vowel even if it reduced, as in (6). At this level, the rule appears to be 'conserve the rightmost vowel.' The same deletion is also observed in nominal derivation.[3]

(6) vula 'read' → vulăttan 'read+Condit./Optative' (UF: -ăttan)
sarlaka 'wide' → sarlakăšĕ 'width' (UF: -ăš+ĕ)

2. Compounding. In LP, compounding is claimed to take place on a separate level from inflection and derivation in some languages. In several LP studies (Mohanen 1984, 1986, Hameed 1985, Shaw 1985) it has been claimed that there are two types of compounding and shown on phonological grounds that two levels of compounding are required in the analysis. These compound types in some instances correspond to the traditional notion of endocentric (headed) and exocentric (non-headed) types. They are often referred to as subcompounds and cocompounds in the LP literature.

Aside from the straightforward claim that these compound types are distinguished in some languages by different phonological rules, it is also claimed that compound types are distinguished semantically. Endo- or subcompounds are sometimes claimed to be compositional and transparent, while exo- or cocompounds are not.

Chuvash shows two major compounding types. One of these is a N N (actually, Substantive Substantive) concatenation in which the elements stand in a qualifier-qualified relationship (Chuvash being a right-headed language).

(8) śul süren
road going (Subst.) → 'traveller'

A second (and typically Turkic) compound type shows the second element apparently inflected for 3Poss Lewis 1967 refers to these (in

Turkish) as 'indefinite izafet' compounds (*izafet* 'annexation' < Arabic); Underhill 1976, also with reference to Turkish, labels them 'possessive compounds.' Example (9) shows such a construction in Chuvash. I shall refer to these as neutrally as possible for now as N N3Poss compounds.[4]

(9) sasă čĕlĕxĕ
voice cord+3Poss → 'vocal cord(s)' < čĕlĕx 'cord'

In an LP approach, a problem arises with this second type of compounding. Specifically, we must ask how we can get what is apparently an agreement inflection on a compounded form that must, all assume, be composed before inflection in the lexicon. For a start, it is to be hoped that we will not have to form such compounds through simple juxtaposition and then crudely mark them as taking the 3Poss inflection at a (later) inflectional level. It would also be preferable not to inflect the second elements and then loop back up the derivational level with the inflected form. It would also be in the interests of simplicity not to have a special level of N N3Poss compounding where the same rules of inflection apply as on the inflectional level. Of course it is possible that N N3Poss compounding occurs on the same level as inflection (see Hameed 1985 for Turkish), but, as I will show, this may not be the case in Chuvash.

Hameed 1985, p. 83, claims for Turkish that one difference between these types is that they are sub- (endocentric) and co- (exocentric) compounds. One type, the uninflected one, is claimed to be semantically transparent since ' ... the head of an endocentric compound provides the thrust of the combined meaning [ozf the whole] while the other is [unheaded] and semantically opaque since the meanings of both items are engaged equally in producing the total, noncompositional meaning.' While it might be hoped that this argument could be made for Chuvash, a look at a number of compounds shows that semantically there is a range of transparency; we are in the presence of a cline, as is so often the case in natural language.

(10)

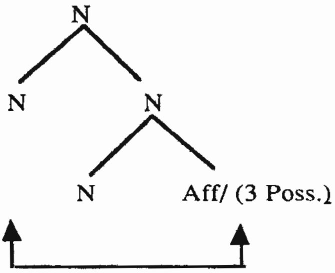
Uninflected, Endocentric Transparent	**Uninflected, Endocentric Opaque**
sasă-čĕvĕ voice sound → 'voice'	xĕvel tuxăś sun rising (N) → 'Orient'
xĕr-ača girl child → 'girl'	mĕn kun great day → 'Spring Solstice'
śum kurăk weed plant → 'weed'	ala śăltăr sieve star → 'Pleiades'
aśa xur male goose → 'gander'	kayăk puś wild-animal head → 'clover'
tisker kayăk robber wild-animal → 'predator'	tɨră kas grain cut → 'ground squirrel'
xura śɨrla black berry → 'blackberry'	sivĕ păx cold crap → 'snowflake(s)'
Inflected, Exocentric Transparent	**Inflected, Exocentric Opaque**
tăval kull-i salt lake-3Poss → 'salt lake'	śer ulm-i land apple-3Poss → 'potato'
kolxoz uy-ĕ kolxoz field-3Poss → 'kolxoz field'	vak kuś-ĕ small eye-3Poss → 'ice hole'
sasă čĕlĕx-ĕ voice band-3Poss → 'vocal cord(s)'	ereśmen kart-i 'spider target-3Poss → 'spider web'

al arman-ĕ hand mill-3Poss → 'handmill'	xurăn sɨrr-i birch berry-3Poss → 'strawberry'
xura śɨrla avr-i black berry bush-3Poss → 'blackberry bush'	sivĕ yen-ĕ cold side-3Poss→ 'north'
suxa puś-ĕ plow head-3Poss → 'plow''	xălxa păx-ĕ ear crap-3Poss → 'earwax'

N N compounds show opacity and idiomaticity just as do N N-3Poss compounds, while some of the N N3Poss compounds are transparent and easily interpretable. Still, an essential grammatical difference remains. N N compounds are always headed, and we can alway see the qualifier-qualified relationship, no matter how metaphorical, idiosyncratic, or idiomatized the meaning has become. Associated with this fact, and maintaining the traditional definition is the ISA relationship as well (Allen 1978), which is to say that the entire compound can be referred to by its head, e.g., 'weed plant' IS A plant, 'robber animal' IS A animal, and so on. This is slightly but significantly different from saying that the head provides the 'thrust of the combined meanings.' Of course, the less transparent compounds and those that are verb derived are not as clear in this respect. N N3Poss compounds sometimes appear to show a qualifier-qualified relationship, but the essential relationship between the elements is different. What is the nature of this difference?

Let us first turn to intepretations of the parallel structure in Turkish. As noted, Hameed 1985 claims that both elements contribute equally to the compositional meaning — a difficult point to defend on semantic grounds. Lewis (undated) 'translates' these constructions as showing agreement. Thus Chuvash *vak kuśĕ* 'ice hole,' would be interpreted as 'small its-eye.' Lewis (1967: 42) attempts to capture the exocentric and often idiosyncratic nature of this compound type in Turkish as follows: ' As a working rule, an undefined izafet group can be turned into intelligible (though not necessarily grammatical) English by a hyphen ...,' thus 'small-eye.' Underhill's 1976 description of this construction as 'possessive compounds' also suggests an agreement relationship (though it may well be that his terminology is purely descriptive). What is implied nonetheless is a relationship between the two Ns of a N N3Poss compound that can be represented as in (11), where the 3Poss on the second N shows agreement with the first, or 'possessing' element.

(11)



Before going on it is important to note that the grammatical relationship expressed in this construction is distinct from a true possessive relationship, as both authors noted above clearly state for Turkish as well. True 'possession' in Chuvash (as in Turkish) is shown formally by cross-referencing the first element is inflected for Gen, and the second for Poss, as shown in (12), though the 3Poss on the second element is omitted in colloquial Chuvash.

(12) kolxoz + ăn uy + ĕ
kolxoz + Gen. field + 3Poss → 'the kolxoz's field'

This is opposed to the compound's meaning 'kolxoz-field,' as given in (10).

It is this possessive relationship that can be said to be truly inflectional, transparent, relational, and even — in the spirit of Anderson 1982 — doable in the syntactic component if necessary. But it is still not clear how we can account for the inflection on the exo-compounds above within the framework of LP. The standard LP model does provide a rather uninteresting way to accomplish this — the Loop. Looping is a controversial aspect of LP that would work as follows in this instance:

on the level of inflection, words are inflected for all cases; then, the required inflected forms loop back up to the previous (compounding) level to be used in exo-compounds. The formal problem here is the power of the loop itself — disputed by many in LP. A second possibilty is to have inflection and exo-compounding on the same level. This is certainly workable, but seems to me to miss the nature of the semantic difference between possessive inflection and the way it is used to form these compounds. Another, particularly unsavory alternative is to have a compounding structure rule of the type $[\,[\mathrm{N}]\,[\mathrm{N}]\ \mathrm{3Poss}]_{\mathrm{N}}$ that would be filled out by having the 3Poss suffix added at the appropriate level.

(13) $[\,[\mathrm{N}]\,[\mathrm{N}]\ \mathrm{3Poss}]_{\mathrm{N}}$

This is workable, but unrevealing, as it suggests that there is no relationship at all between the two elements.

3. A Solution. I propose that the apparent 3Poss affix used in Chuvash cocompounds is a special type of affix, distinct semantically from the other posessive markers; it is this difference in its semantic content that allows it to function as it does. Its semantic function is, broadly speaking, to act as a definitivizer. The distinction between two kinds of application of this affix was noted in a historical framework in Benzing 1959.[5] In the framework I am employing here, Chuvash has two lexical levels. Uninflected stems are derived on Level 1; compounds and productive inflection occur on the second level.

The level of stem derivation is characterized by different morphophonemics than the compounding/inflection level, as has been shown in (2) - (6). In fact, this level's morphophonemics is strikingly impoverished when compared to that of the stem-inflection interactions at Level 2. In other words, we see more morphophonemic action as we approach the right margin of the Chuvash (and, broadly speaking, Turkic) word. This probably reflects a historically later agglutination of these affixes and hence a less phonologically settled and grammaticized layer of affixing.

I propose that the 3Poss is applied as a derivational suffix (henceforth 3Def.) at Level 1 to create cocompounds. A second, homophonous affix, the real 3Poss applies inflectionally at Level 2. There are several facts that allow us to interpret 3Def as derivational in nature.

(a) 3Def does not alter the syntactic relations between the two members of the compound, but the semantic relations. Definiteness is not inherently a syntactic property, but a semantic (conceptual) one. Derivational affixes typically alter the conceptual meaning of bases to which they are attached. Thus, the 3Def in Chuvash may apply to any Substantive with the resultant change in the meaning to [+specific], 'one who ...,' and so on. For example, alone and outside of compounds, *kuś* means 'eye,' or 'the eye,' while *kuśĕ* means 'the/some (specific) eye.' It is interesting to note that specifying is also carried out by another so-called 'inflectional' affix in Chuvash, the plural. The phrase *pilĕk pürt* means '(any) five houses'; the phrase *pilĕk pürtsem* means 'five specific houses'; just as the form *pürtsem* means 'a number of (specific) houses,' not just 'houses.'

(b) The 3Def is employed contrastively with Postpositions to mark them as more specific. Thus *xĕr* 'edge' vs. *xĕrri* 'front, side' ← *xĕră+i* 'edge + 3Def.' Furthermore, the resulting construction can be inflected for case and then possession, as shown in (14) (the *n* of the 3Def and 3Poss is in the UF and only surfaces when an inflectional morpheme follows; recall here that the forms of the 3Def./3Poss is *-ĕ* after stems that end in a C and *-i* after stems that end in a V).

(14) xĕr 'edge' → xĕrri 'front side' < xĕrr + i 'edge+3Def.'
pat 'side'
pat + ĕ '(specific) side'
patĕnči 'on/at its side' < pat + ĕn + ra + i
side +3Def.+Loc.+3Poss

(c) Once specified, the postpositional nominals are subject to regular derivation, as would be expected of other derived forms. In other words, this particular suffix is a word-builder and occurs inside other word-building affixes, as (15) shows.

(15) { { { věs + ĕ }$_N$ { xĕrr + i }$_N$ } sĕr }$_{Af}$ }$_N$
{ { { end+3Def }$_N$ { edge+3Def. }$_N$ } less }$_{Af}$ }$_N$ 'infinity'

(d) The 3Def is used after case forms to specify cased Nominals, as in (16), thus providing some 'derivation through inflection' at the Compounding/Inflectional level. Here, we may be forced to accept limited looping.

(16)

anat		'lower (Chuvashia)'
anatra	anat + ra lower + Loc.	'in lower Chuvashia'
anatri	anat + ra + i lower + Loc. + 3Def.	'person from lower Chuvashia'
puśri	puś + ra + i head + Loc. + 3Def.	'a thing on the head'

Taken together, the evidence accumulates to suggest that this quasi-derivational element creates substantival forms that are semantically 'specific.' Its effects are created on a level at which it can freely interact with compounding, case inflection, and possessive inflection (I return to the interesting problem represented in example (15) below).

Thus, what is special about the 3Def is that it does not act as a possessive inflection, but as a definitivizer. Its (ad hoc) feature content is shown in (17).

(17)
$$\left\{\begin{array}{ll} + & \text{definite} \\ \emptyset & \text{person} \\ \emptyset & \text{number} \end{array}\right\}$$

Here 'Ø' means 'lacking.' In contrast, other possessive affixes have the feature content of (18).

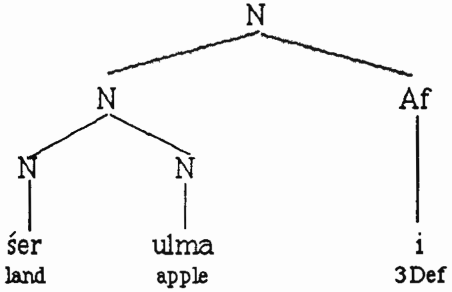
(18) $$\left\{\begin{array}{ll} + & \text{definite} \\ \varnothing & \text{person p} \\ \varnothing & \text{number n} \end{array}\right\} \text{(where p = 1, 2, or 3 and n = S or Pl)}$$

All possessive suffixes are semantic specifiers. All indicate that the possessed N is specific or definite simply by virture of its being possessed. But 3Def is a specifier devoid of person/number content — it is a 'pure' definitivizer.

Whatever the implications of this analysis for other languages, this proposal does not appear to be coincidental in a language family where there is no gender marking. Chuvash *vun* (= Turkish *o)* means 'he, she, it' as well as serving as the demonstrative 'that.' Furthermore, 3S and 3PL possession is marked with the same form; no number is shown in 3Poss

Returning to subcompounds, we can see that 3Def does not relate one N to another by agreement as shown in example (11), but characterizes the compound by semantically marking the entire construction as definite, that is, as semantically unitary. The representation of the cocompounds is thus as in (17).

(19)



Here, 3Def indicates a marked compound that is less likely to show compositional meaning. As we noted, despite the occasional opacity of unmarked N N compounds as opposed to the marked types, the former still hold generally to the ISA relationship (they are clearly headed) while the latter do not. The salt lake of 'salt lake+3Def.' in indeed a lake, but it is a specific and thus definite kind of lake; the kolxoz field of 'kolxoz field+3Def.' is a specific kind of field. But the 'eye' of 'small eye+3Def' = 'ice hole' is not an eye at all, and the 'apple' of 'land apple+3Def.' is really a potato.

Finally, as already noted, the apparent occurence of Level 1 derivation outside Level 2 compounding and Inflection in example (15) may have to be accounted for by still permitting a limited amount of looping. But the looping of compounds or even sentential phrases is attested — or at least claimed — in a number of LP studies of diverse languages (English, Dakota, Malayalam, Welsh) to account for productive curiosities like 'black eyelessness.'

NOTES

[1] Transcription follows the North American standard with the following additions: *ă* and *ĕ* represent low back and front reduced vowels, and *ś* , a voiceless palatal fricative.

[2] A rule of dissimilation accounts for the change of *r* to *t* ; an assimilation rule accounts for the change of *m* to *n*. Both rules are independenly motivated. See Kreuger 1961 for more data.

[3] There are some peculiar exceptions that I have not resolved.

[4] The 3Poss form has two front vowel allophones whose alternation is predictable; the vowel is always *-i* when affixed to a vowel-final stem (and replaces the stem final vowel), and *-ĕ* when affixed to a consonant final stem. Certain forms, notably those which end in a reduced vowel, show apparent gemination in their possessive forms and certain case forms (cf. Dobrovolsky 1984 for a now outdated analysis.)

[5] Thanks go to Éva Czató for recalling this to my attention.

REFERENCES

Allen, M.R. 1980. 'Semantic and Phonological Consequences of Boundaries: A Morphological Analysis of Compounds.' In Aronoff, M. and M. L. Kean (eds.), *Juncture*, 9-27. Amna Libri.

Anderson, Stephen. 1982. 'Where's Morphology?' *Linguistic Inquiry* 13: 571-612.

Ašmarin, N.I. 1968. *Thesaurus Linguae Tschuvashorum*, II. Reprint: Kazan' (1929). Indiana University Pulications Uralic and Altaic Series, 70/2. Bloomington: Indiana University Press.

Benzing, Johannes. 1943. *Kleine Einführung in die Tschuwaschische Sprache*. Berlin: Otto Stollberg.

____. 1959. 'Das Tschuwaschische.' In *Philologiae Turcicae Fundamenta* , I: 695-751. Wiesbaden.

Dobrovolsky, Michael. 1984. 'Abstract and Concrete in Chuvash Morphophonemics.' *Folia Slavica* 7: 120-134.

_____. 1989. 'Towards a Lexical Phononogy of Chuvash.' In Aronson, Howard I. (ed.),*The Non-Slavic Languages of the USSR: Linguistic Studies* , 80-92. Chicago: Chicago Linguistic Society.

Hameed, J. 1985. 'Lexical Phonology and Morphology of Modern Standard Turkish.' *Cahiers linguistiques d'Ottawa* , 14: 71-95.

Kaisse, Ellen and Patrica A. Shaw. 1985. 'On the theory of Lexical Phonology.' *Phonology Yearbook* 2: 1-30.

Kiparsky, Paul. 1982. 'From Cyclic Phonology to Lexical Phonology.' In van der Hulst, Harry, and Norval Smith (eds.), *The Structure of Phonological Representations*, 131-176. Cinnaminson, N.J.: Foris.

Kreuger, John. 1961. *Chuvash Manual.* Indiana University Publications in Linguistics, Uralic and Altaic Series 7. Bloomington: Indiana University Press.

Lewis, G.L. Undated. *Teach Yourself Turkish.* London: English Universities Press.

_____. 1967. *Turkish Grammar.* London: Oxford University Press.

Mohanen K.P. 1984. *Lexical Phonology.* Bloomington: Indiana University Linguistics Club.

Mohanen K.P. 1986. *The Theory of Lexical Phonology.* Boston: D. Reidel.

Paasonen, H. 1908 (1974). *Tschuwaschisches Wörterverzeichnis.* Edited by A. Róna-Tas. [Studia Uralo-Altaica IV.] Szeged: Universitas Szegediensis, Amsterdam: John Benjamins.

Shaw, P. A. 1985. 'Modularization and Substantive Constraints in Dakota Lexical Phonology.' *Phonology Yearbook* 2: 171-200.

Scalise, Sergio. 1986. *Generative Morphology.* Dordrecht, Holland, Riverton, USA: Foris.

Underhill, Robert. 1976. *Turkish Grammar.* Cambridge, Mass.: MIT Press.

Vikár, László and Gábor Bereczki. 1979. *Chuvash Foksongs.* Budapest: Akadémiai Kiadó.

LAK SUBSTANTIVAL DECLENSION: 40 CASES OR 50?

Victor A. Friedman
University of North Carolina, Chapel Hill

According to the *Guiness Book of World Records,* Tabasaran holds the record for the largest number of cases. Second only to Tabasaran, Lak has the world's next largest number of cases, as well as the next smallest number of speakers among the five Daghestanian literary languages. The student of Lak, however, will be confronted not merely by a large number of cases, but also by conflicting figures and definitions. Thus, for example, Uslar (1890:16) sets the number at "about 50" and actually gives examples of 48 (50 if the adverbials *-nu* and *-nā*— are counted). Dirr (1928:238-41) uses Uslar's 48. Hjelmslev (1935:159-83) also sets the number at 48, but his inventory differs from both Uslar's and Dirr's. From Bouda's (1949:38-53) account one can deduce 45 cases, but again some of them differ from those cited in previous works. Burçulaje (1986:80, 117) states "42 (sometimes 41)," while Žirkov (1955:35-37) gives only 40 forms in his "complete paradigm." Both these authors cite desinences not cited in anyone else's nor in each other's work. Murkelinskij (1971:128-134, 1980:492-93) leaves out one of Žirkov's cases (the comparative *-jar)* to get 39. These last three authors also indicate that certain of the case desinences they cite are no longer in use. In addition to disagreeing on the inventory of Lak cases, authors also disagree on the classification of certain desinences, i.e., as primary or grammatical, secondary or local, and sometimes tertiary or adverbial. In this paper I hope to bring some order to this confusing state of affairs.

I shall begin with a few basic facts of Lak nominal morphology. There is a primary distinction in Lak between the nominative stem and the oblique stem. The oblique stem is often but not always characterized by an additional suffix. Vocalic and/or consonantal alternations or suppletion also occur. Some typical examples are illustrated in Table One.

TABLE ONE

meaning	*nominative [class]*	*genitive*
eagle	barzu [3]	barzu-l
game	.turḳu [4]	.turḳulu-l
sister	ssu [3]	ssi-l
place	ḳanu [3]	ḳanttu-l
head	baḳ [4]	baḳra-l
heart	daḳ [4]	daḳni-l
side	čul [3]	čulu-l
field, steppe	čul [3]	čuldanu-l
stable	ppal [3]	ppalni-l
wool[sheep,camel]	ppal [4]	ppalu-l
wolf	barç [3]	burçi-l
back	barḫ [3]	burḫa-l
word	maq [3]	muqqu-l
mother	ninu [2]	nitti-l
sun	barz [3]	zuru-l

Lak also has an important distinction between personal (first and second) and non-personal (third) pronouns, with suppletion in the oblique personal singular and class marking in the non-personal as well as in the emphatic/reflexive personal. Table Two gives the nominative and genitive of the personal pronouns and a typical non-personal one. Table Three gives two singular emphatic/reflexive pronouns — first and third singular — in enough cases to illustrate all the variants of the word-internal class markers. Also the the initial and plural markers are listed below the chart.

TABLE TWO

	singular		plural	
	nom	gen.	nom.	gen.
first person	na	ttu-l	žu	žu-l
second person	ina	vi-l	zu	zu-l
third: class 1	ga	gana-l	gaj	gajnna-l
2		gani-l		
3				gajnnu-l
4				

TABLE THREE

CASE	CLASS 1	3	2,4	1	3	2,4
nom	na-v-a	na-v-a	na-r-a	cu-v-a	cu-pp-a	cu-rd-a
gen	ttul-ø-a	ttul-v-a	ttul-l-a	cal-ø-a	cil-ø-a	cil-l-a
dat	ttun-ø-a	ttun-m-a	ttun-n-a	can-ø-a	cin-m-a	cin-n-a
supr	ttuj-v-a	ttuj-v-a	ttuj-r-a	caj-v-a	cij-v-a	cij-r-a
	I myself			oneself		

other class markers:

initial: 1 = ø, 3 = b-, 2,4 = d-   plural: 1,2,3 = 3sg, 4 = 4sg.

With these basic facts in mind, I shall now attempt to synthesize as coherently as possible the various accounts of case in Lak cited above. For the time being, I shall follow all the scholars from Uslar to Burçulaje in taking as given the meaning of the term *case*. The first question is whether or not to ascribe an ergative case to Lak. As is illustrated by the example sentences below, that which corresponds to the English subject can be in the nominative, genitive, ablative, or dative case, while that which corresponds to the English object is in the nominative. The subject of an intransitive verb is nominative (1). The transitive verb (here *axxan* 'to sell') always agrees with the object in class (here *b-[v]* marks *čvu* [3] 'horse'). In an ordinary simple transitive sentence, however, a personal pronoun will be nominative while anything else will be genitive (2 & 3). The nominative can control person agreement, but the genitive cannot (1, 2 & 4). Note that if the tense form is analytic, non-personal subjects have the option of being in the nominative and controling the auxiliary (4b). According to Kibrik (1978:9), the different agreement patterns are used to signal degree of commitment to the statement (3 & 4), while the ablative can be used to signal unintentional performance of the action (5). The ablative is also required by verbs meaning 'be able' (7), and the dative is required by so-called 'affective' verbs, e.g., 'want' and 'see' (6).

1a. na naj ura	I am coming
1b. ga naj ur	He is coming
2a. na baxxara čvu	I sell the horse
2b. ganal baxxaj čvu	He sells the horse

3a. na bavxxunu bur čvu	Apparently I sold the horse
3b. ganal bavxxunu bur čvu	Apparently he sold the horse
4a. na bavxxunu ura čvu	I have sold the horse
4b. ga bavxxunu ur čvu	He has sold the horse
5a. tušša bavxxunu bur čvu	I accidentally sold the horse
5b. ganašša bavxxunu bur čvu	He accidentally sold the horse
6a. ttun va čvu baxxan ččaj bur	I want to sell this horse
6b. ganan va čvu baxxan ččaj bur	He wants to sell this horse
7a. ttušša va čvu baxxan bjuqlaj bur	I can sell this horse
7b. ganašša va čvu baxxan bjuqlaj bur	He can sell this horse

Kibrik (1978:3) considers the ergative to be a separate case that is realized on the surface as either nominative or genitive depending on the lexical characteristics of the noun. As seen in 4 however, the semantics and shape of the tense form can also affect case choice. Kibrik's formulation obscures the fact that like many of the world's languages, Lak has a type of split ergative structure that conforms to the feature hierarchy described by Silverstein (1976).[1] We thus agree with Burč̣ulaje (1986:84) that Lak cannot be considered to have an independent ergative case.

The next question concerns the distinction between primary (grammatical, abstract), secondary (local, concrete) and tertiary (adverbial) cases. None of the Lak cases are purely grammatical in the sense that Mel′čuk (1986:84) uses for the Georgian ergative. All Lak cases carry some sort of semantic valence beyond that of syntactic role in the clause. Some do appear to be purely local or adverbial, but others are more difficult to classify. All scholars are agreed on at least three grammatical cases: nominative *-ø*, genitive *-l*, and dative *-n*. Moreover, there is a series of agglutinatively formed local cases that everyone agrees form a coherent set, although the exact number of actually occurring and semantically distinct forms varies from one account to another, with no account attesting all forms and meanings.

The local cases are formed by adding one of six morphemes to the oblique stem and one of seven morphemes (including zero) to that suffix, bringing the theoretically possible total to 42, to which we must add the fact that Bouda (1949:46) distinguishes *-v* from *-vu*, while all others treat *-v* as an ellided variant of *-vu*. Bouda's claim pushes the possible total of these local cases to 43. Table Four gives the six non-terminal local case desinences, the seven terminal local case desinences, and all the possible combinations, including class agreement for *-n!aj* (where *!* stands for the position of the class marker).[2]

TABLE FOUR

suffix	*meaning*	*label* *(After Mel'čuk 1986 and Hjelmslev 1935)*
-v[u]	in	in-
-j[3]	on	super-
-x	behind	post-
-lu	below	sub-
č̣a	near	apud-
c̣	at	ad-
-Ø	location	-essive
-a	from	-ablative
-atu	away from	-elative
-ux	through/across	-prolative
-xc̣in	through/via	-vialative
-un	to	-lative
-un!aj	toward	-directive

v					
vu	j	lu	x	č̣a	c̣
va	ja	la	xa	č̣aa	c̣a
vatu	jatu	latu	xatu	č̣atu	c̣atu
vux	jx	lux	xux	č̣ax	c̣ux
vuxc̣in	jxc̣in	luxc̣in	xuxc̣in	č̣axc̣in	c̣uxc̣in
vun	jn	lun	xun	č̣an	c̣un
vunaj	jnaj	lunaj	xunaj	č̣anaj	c̣unaj
vunnaj	jnnaj	lunnaj	xunnaj	č̣annaj	c̣unnaj
vunmaj	jnmaj	lunmaj	xunmaj	č̣anmaj	c̣unmaj

Before broaching the question of case vs clitic postposition as well as the question of tertiary or additional local cases, let us put this vast assortment of forms into perspective by defining who recognizes what as an independent case form.

As was already indicated, only Bouda (1949:46) separates *-v* from *-vu*, while all other authors treat *-v* as a variant of *-vu* (e.g., Žirkov 1955:41). Bouda claims that *-v* is lative while *-vu* is essive, giving examples such as the following:

8a.	dušru ššinav naj bija	the girls go for [to fetch] water
8b.	bjarav bahnu	having fallen in a pond
9a.	ššinavu labivḳun	hidden in the water
9b.	Kavkaznavu dja″vi bija	There was war in the Caucasus

Consider, however, the following examples, which clearly show *-v* as essive and *-vu* as lative:

10. Ttun ččaj bur, ... ttula šjaravu C̣aday o′rmu butlan. (Ḫamzatov 1972:19)
I want ... to spend my life in my own aul, in C̣ada.

11. Šikku, Moskavliv, dukra dan vilaxxay c̣u qqadišajssa xxanssar, (Ḫamzatov 1972:19)
Here, in Moscow, it appears fires aren't lit in the hearth to cook.

12. Bivxxun bur ca ḫajvan, bivx′unu bur darvagravu va šanijn buvtun bur jalttu birhangu lirčunu. (Xalilov 1976:226)
They slaughtered an animal, put it into a sack and put it on the bed, having thrown clothes on top.

13. Na, ca zurujssa otpuskgu lavsun, buttal šjaravun, C̣aday lavgssijav. (Ḫamzatov 1972:18)
I took a month's holiday and went to my native aul to C̣ada.

It thus does not appear that *-v* and *-vu* are distinguished on the basis of an inherent meaning. It is more likely that there is free variation with many nouns, and a tendency to favor one or the other ending in certain nouns or types of syntagma, e.g., compound verbs.[4]

The pre-1950 authors, i.e., Uslar, Dirr, Hjelmslev and Bouda, all separate *-a* and *-atu*, whereas the post-1950 authors, i.e., Žirkov (1955:38), Murkelinskij, Burč̣ulaje, as well as Gajdarova (1977), all treat *-atu* as a facultative variant of *-a*. Xajdakov (1966:15) states quite explicitly: "The affix *-tu* performs the same function as *-a* and never occurs without it. Its use is facultative." With the *apud*-formant *-č̣a*, only Hjelmslev (1935:167) opposes an apudablative *-č̣aa* to an apudelative *-č̣atu*. All other authors either clearly imply or explicitly state that only *-č̣atu* occurs (e.g., Uslar, 1980:45, Gajdarova 1970:315, Burč̣ulaje 1986:81). Given that Hjelmslev's account is based entirely on secondary sources and that this is the one case for which he adduces no textual examples, it must be dismissed as the apparent result of Hjelmslev's wish to describe Lak in terms of a perfectly symmetrical local case system. The pre-1950 authors all follow Uslar (1890:37) in claiming that although *-a* and *-atu* are often synonymous, *-atu* nonetheless adds a nuance of distance or removal. Bouda (1949:47) describes *-a* as more ablative and *-atu* as more elative. He suggests the relationship of *-a* to *-atu* is comparable to that of the -lative *-un* and the directive *-un!aj*.

The following example is typical:

14. Č̣iravatu šatta buvkkunni. (Murkelinskij 1971:130-33)
The snake came out from inside of the wall'

but also:

15. Šatta buvkkunni murx'irala. (Murkelinskij 1971:130-33)
The snake crawled out from under the tree'

Here, however, are some examples with *-atu* where the sense of 'distance' or 'removal' is difficult if not impossible to identify:

16. Vin čatu ḳulli. (Xalilov 1976:228)
How (From where) do you know?[5]

17. Musijatu usttarnal ḳissa buvna. (Gajdarova 1977:278)
The mastercraftsman made the ring out of gold.

18. O′nnassa č̣umal ga ṭannujatu žura-žurassa qqusri, kkačči-ččitri jalagu cajmi ssix′ri dullan iḳajva. (Gajdarova 1977:278)
He made various spoons and toys and other things out of wood in his free time.

Although *-un* and *-un!aj* also have similar, sometimes near-synonymous, uses, there are also examples where they are clearly not interchangeable:

19. Na tanajn zakkana ššun bav (Xajdakov and Žirkov 1962:327)
I hit [onto] him with my fist.

20. mağujnaj uruglaj ur. (Murkelinskij 1971:131)
He is looking towards the ceiling.

21. ... o′vča vila dusturajn va mallanajn pulav bukan, ... (Xalilov 1976:205)
... invite [onto] your friends and the mullah to eat pilaf, ...

To the best of my knowledge, such examples do not exist for *-a* and *-atu*.

Of the authors whose works were available to me, only Burč̣ulaje identifies the series based on the terminal desinence *-xč̣in,* which he credits Topuria(1940) with having first observed. He does not make clear how *-xč̣in* differs from *-[u]x*. So far I have found only two examples, both of them super-:

22. Zu zula dua′rttajxč̣in hattava qqaliqaj ara, paraq̇at ara,—kunu. (Xalilov 1976:216))

By means of your prayers make it so he doesn't come running back from the grave, make him rest in peace—she said.

23. Ç̌alaj buriv zuva ussurval bjavq̇laj bivḳšivu va kulpatravussa dusšivu žujxç̌in dušivu—kunu. (Xalilov 1976:225)
'Do y'all see that you brothers are mistaken [you brothers' being mistaken] and that the family friendship is through us [the family friendship's being through us]—she said.'

While both these examples are non-local, i.e., non-literal, such usage also occurs with the superprolative *-jx,* although I found only one example as opposed to twelve with local meanings.

24. Ina ura, ttul muqqurtijx uvkkun, qqalmaqqallaj. (Xalilov 1976:225)
Through my words, you are disgraced.

It is not possible to comment on this case further at this time.

With regard to the remaining cases, while the post-1950 authors state that some of them are obsolete, they are all attested as distinct. There are clearly differences among dialects and between spoken and written usage in this regard. Thus while the system of local cases in Lak is undergoing change (Burçulaje 1986:82), at this point in time the remaining cases must all be counted in a complete inventory.[6]

We must now return to the question of whether all these local cases are really cases or clitic postpositions. Lak has a number of non-clitic postpositions. Of this number, five correspond more or less closely to the six primary local case desinences as shown in Table Five:

TABLE FIVE

case	meaning		postposition	meaning
-vu	in-	in	viv	inside
-j	super-	on	jalu	above
-x	post-	behind	maq	behind
-lu	sub-	below under	luv	below

-č̣a	apud-	near	č̣arav	around, at, by[7]
-c̣	ad-	at	"	"

Although these local cases are clearly postpositional in origin, there are several arguments in favor of analyzing them as synchronic cases.

1. Secondary local cases can be followed by postpositions and postpositions can govern more than one case:

25. qqatra-l maq
behind/at the back of the house [pl]
26. Gani-j-a maq ars vaj šanmagu zaraja maqunaj x′unu ur. (Xalilov 1976:212)
After that [superelative] the son swore off those three things.

2. Constructions with postpositions and their semantically related cases have different meanings:

27a. qqatlu-vu ‘in the house’
27b. qqatlul viv ‘inside the house’

3. Secondary local cases can have non-local meanings, postpositions are restricted to concrete meanings:

28. ... o′vča vila dustura-j-n va mallana-j-n (≠ dustural va mallanal jalun) pulav bukan, ...(Xalilov 1976:205)
... invite [onto] your friends and the mullah to eat pilaf, ...

29. durnu dur dukra hutru - ttiliḳra-j-a (≠ ttiliḳral jala). (Xalilov 1976:219) ‘
They made food from the pluck [liver & lungs].

30. Tana-j-a (≠tanal jala) rasxat lasi. (Murkelinskij 1971:131)
Get permission from him

31. qqatlu-x (≠qqatlul maq) cij dulav? (Murkelinskij 1971:132)
What did you pay for the house?

32. tana-x-a (≠ tanal maqa) uvkkun havav? (Murkelinskij 1971:133)
Are you leaving because of him?

33. Nittil buruvxxussa huqra-v-a (≠ huqral viva) Tamaril jaṭjulssa huqa buvç̌una. (Gajdarova 1977:277)
From the dresses sewn by mother Tamar chose a red dress.

34. Vana lax′x′u radiolu-vu-x (≠ radiolul vivux) buslaj bija dušmannal ğužum bjajššin bunni ṭij. (Gajdarova 1977:279)
Yesterday they announced on the radio that the enemy's assault had been repelled.

35. A′q̇dulu-ç-a šajar čanna-ç-a (≠ a′qḫdulul/čannal ç̌arava) x′unu qinssar. (Gajdarova 1977:280)
Than deprived of reason deprived of sight (lit. light) is better.

4. Some postpositions are formally very different from the semantically similar case:

36a. qqatlu-l maq 'behind/at the back of the house [pl]'
36b. qqatlu-x 'behind/at the rear of the house'

5. Postpositions can take secondary case suffixes that differ significantly from the ones used in nominal declension:

TABLE SIX

-essive	-lu	luv	-j	jaluv
-prolative	-lux	lulttu	-jx	jalttu

6. Postpositions can govern conjoined genitives, case forms are repeated each time:

37. ... ššjalmaḫral *va* ṭajlamunil djaniv*gu* xxjuva ḳissa bišinssa ḳanu bussar. (Xalilov 1976:211)
... and the distance between truth *and* falsehood is five fingers.

38. ... o′vča vila dustura-j-n va mallana-j-n pulav bukan, ...(Xalilov 1976:205)
... invite [onto] your friends and the mullah to eat pilaf, ...

7. Class markers are bound morphemes, not clitics, and case desinences occur before class markers in emphatic pronouns. Clitics and postpositions follow the class marker.

39. Ga na na-va-ra (Ḫamzatov 1972:21)
It is I myself (I-class-am)

40. Ina-va-gu (Xalilov 1976:219)
You yourself, too (you-class-and)

41. ca-l-ø-a x′x′ič̣ ‘in front of him himself’ (him-gen-class + postposition) (Xalilov 1976:204)

Similarly:

ttu-č̣an-ø-a - ttu-č̣an-m-a - ttu-č̣an-n-a ‘over by me myself’
ca-č̣an-ø-a - ca-č̣an-m-a - ca-č̣an-n-a. ‘over by him himself’

8. Postpositions can occur separated from the words they govern by other parts of the sentence:

42. *Ttul* x′x′ič̣ murx′, maq č̣ira bur. (Murkelenskij 1983:184)
Before *me* is a tree, behind a wall.

The remaining cases are summarized in Table Seven. I have used the author’s term wherever possible. When the author used a preposition instead of a term, I have supplied an appropriate term:

TABLE SEVEN

suffix	*Murk71*[8]	*Dirr*	*Žirkov55*	*Burč̣86*	*Uslar90*	*Bouda49*
-x′	adessiv	poss′v	poss′v	poss′v	poss′v	adess′v
-x′x′un	dat-2	admot′v	dat-2	adlat′v	adlat′v	allat′v
-ššа	elative	ablative	abl′v	elative	elative	ablative
-ššal	commit′v	comit′v	comit′v	comit′v	comit′v	comit′v
-jnu	instr′l	instr′l	caus′v	inst′l	inst′l	caus′l
-xlu[nu]	motivat′v	—	motivat′v	motivat′v	—	—
-[ni]jar	=adv	comp′v	comp′v	comp′v	comp′v	=abl+-r
-[u]kssa	=adv	similat′v	≠case	—	equat′v	=učin
-[u]kun!a	=adv	similat′v	≠case	—	similat′v	=učin
-j	=nom[86] voc[126]	voc	≠case	—	=superes′v	≠case

Two questions are involved here. The first is which if any of these are cases? The second is whether they are to be classed as local, grammatical, or something else. Let us begin by dismissing Hjelmslev's (1935:101) claim that *-gu* forms a separate, similative, case. He cites the following sentence:

43. Ussu-gu ssu̱-kssa qinssa uri
Brother is as good as sister.

In fact, however, *-gu* is a clitic whose basic meaning is 'and/even,' not a case suffix. It can attach to any case and to any part of speech. Examples (37) and (40) show it attached to a postposition and following the class marker on an emphatic/reflexive pronoun. In example (44) it is attached to a dative and in (45) it clearly has no comparative/similative meaning and is moreover attached to oblique case suffixes:

44. Na ussi-n-ukssa ssi-n-*gu* dulav. (Žirkov 1955:44)
As much to brother *also* to sister I gave = I gave as much to sister as to brother.

45. Daḳ čun dullan ivḳun ur, ussi-j-atu-*gu* arx uvcun, cavaj vixšala daqqassa dusta-l-*gu* buvnu,... (Xalilov 1976:225)
He became arrogant *and* turned away from his brother, *and* made some faithless friends...

The suffixes *-[u]kssa*, 'as, like (quantity),' *-[u]kun!a* 'as, like (quality)' both attach to the unaltered nominative. Moreover, as seen in example (44), they can also attach to oblique cases (*ussinukssa*), as well as to other parts of speech, e.g., the adjective in (46), and *kuna* can even stand alone as in (47):

46. Ganal a′rkinssa-kssa arcu darṭun, çaxxučunax′x′un dullunu, kumag buvnu bur burž laqan. (Xalilov 1976:223)
He gathered as much money as necessary, gave it to his neighbor, and helped him to pay off his debt.

47. Na ta kuna jarxnu aqqara. (Murkelinskij 1971:248)
I am not as far away as he.

We thus side with Žirkov and the other scholars who do not consider these case suffixes.

The vocative is always problematic. There is the primary question of whether the vocative is a case at all in any language. In Lak, it only has a distinct form if the noun ends in a vowel other than *-u*, in which case the suffix is *-j* added to the nominative. While functional arguments against the vocative as a case are not language specific, we can add here the specific fact for Lak that cases are added to an oblique stem and are formed agglutinatively for all nouns whereas the vocative meets neither of these criteria[9]

One other troublesome suffix is the comparative in *-jar*. This suffix attaches to the oblique stem, e.g *ttujar* 'than me,' *vijar* 'than you' and in the emphatic/reflexive it comes before the class marker: *vijarva* 'than you yourself.' The apparently completely synonymous *-nijar*, however, attaches to the nominative stem. Bouda's analysis of it as the superelative-*j-a* plus *-r* (class marker?), while tempting, leaves the problem of *-r* unresolved. Another point that Bouda makes that seems stronger is the fact that *-jar* has functions that could be interpreted as broader than that of a strict comparative:

48a. Camur xxullijx x′unu, gana-jar x′x′iç̌un uvkkun, (Xalilov 1976:213)

Taking another road, he came out [having come out] in front of him [ahead of him]

On the other hand, consider also this use with a gerund:

49. A′q̇dulça šajar čannaça x′unu qinssar. (Gajdarova 1977:280)
It is better to be deprived of light [=sight] than of reason (Reason-ad/elative being-than light-ad/elative having-become good-is.

Murkelinskij (1971:248) treats it as an adverbial suffix like *-kun!a*, although this does not explain away the oblique stem. The attachment to non-nominal forms as in (48) favors Murkelinskij's view. If we take a strictly formal view, *-jar* must be viewed as a case suffix because it attaches directly to the oblique stem of any noun.

The remaining cases all attach unambiguously to oblique nominal stems. Of these *-jnu* and *-xlu[nu]* are more "adverbial" while the other four form a set similar in some respects to the local series but much more highly grammaticalized. Here are some typical examples of usage of each case:

Possessive/adressive -X':

50. Tana-x' ttul lu bussar. (Murkelinskij 1971:129)
He has my book.

Note especially the ergative pattern with verbs of speaking:

51. Dušnil butta-x′ kunu bur: — Ina lavgun xanna-x′ uča: (Xalilov 1976:211)
Daughter to father said: "You go and say to the khan:

52. Žula 15 o′l bur, amma canni-x′-ra-gu vakssa naḫussa naķ daqqar. (Xalilov 1976:210)
We have 15 cows, but not one of them has such tasty milk.

Dative-2/admotive -X´X´UN:

53. Muni-x´x´un qqulajssa tahargu dirimu dur. (Xalilov 1976:.204)
A convenient situation occurred for this.

54. Nadiršahnan ččan bivḳun bur lakral bilajat cala ḳuntti-x´x´un lasun. (Xalilov 1976:207)
Nadir Shah wanted to take the Lak country into his hands.

55. Lavsunu ččarulliv ca-canna-x´x´un ca-ca bullunu bur cala arsnan va ussil arsnan. (Xalilov 1976:225)
Taking the kidneys he gave one apiece to each of them—one to his own son and one to his brother's son.

Ablative/involuntative/possibiliative -ŠŠA:

56. Na tana-šša lu lasav. (Murkelinskij 1971:130)
I got a book from him.

57. X̣uxči-šša nic bivḳunni. (Murkelinskij 1971:129)
The shepherd killed the ox by accident.

58. Va davu tana-šša qqurtal dan qqašajssar (Murkelinskij 1971:129)
This job can't be finished by him.

59. Tana-šša ca ḳanttaj ac̣an qqašaj. (Murkelinskij 1971:129)
He can't stand in one place.

60. Tu-šša ḫisav dan bjuqunni. (Murkelinskij 1971:130)
I can solve the problem.

Comitative -ŠŠAL:

61. Ussu ssi-ššal bazalluvun lavgunni. (Murkelinskij 1971:130)
Brother with sister went to the bazaar.

62. Ğaraldanu-ššal miḳgu bija. (Murkelinskij 1971:130)
The town, too, was with rain.

Instrumental/causative -JNU:[10]

63. Vil kumarga-jnu ttušša zanan bjuqanssar. (Murkelinskij 1971:130)
With your help I can walk.

64. Tana-jnu davu qqurtal x'unssa. (Murkelinskij 1971:130)
Through him the job will be completed'

Motivative -XLU[NU]:

65. Ttul ussu buttal ḳanttu-xlu o' eq̇i buṭlaj ussija. (Murkelinskij 1971:130)
My brother shed blood for the fatherland.

66. Ina tana-xlu arcu dula. (Murkelinskij 1971:130)
Pay the money for him.

67. Ta ttu-xlu uvkunni. (Murkelinskij 1971:130)
He acted for me.

As can be seen from the foregoing examples, if we eliminate those suffixes that attach to the nominative stem and leave to one side the problematic *-jar*, the remaining six suffixes divide into two types which could be called pseudo-local and, following Burç̣ulaje (1986), postpositional. The so-called possessive, dative-2, and ablative clearly are more grammatical than local in their functions. The comitative, while clearly etymologically related to these first three, does not appear to have the same types of purely grammatical functions. The last two cases both look rather adverbial; in fact, *-nu* is the derivational suffix for forming adverbs. Moreover, the functions of *-xlu* can also be performed by the postposition *çani*. Still, their syntagmatic features, i.e., attachment to the oblique stem of any noun, as well as their semantics appear to be declensional. Žirkov's (1955) separation of the possessive/addressive and dative-2 into a seventh, defective series of local cases on a par with the more strictly local cases does not seem justified in view of their primary functions to indicate possession and indirect objects of verbs of speaking and giving. The so-called ablative has important grammatical functions, and the comitative can be placed with the other three fairly easily. Mukelinskij's treatment is reminiscent of Žirkov's, but Gajdarova and Burč̣ulaje clearly do not consider any of these cases part of the strictly local case series. Burč̣ulaje's description of all these cases as "postpositional" appeares to be too extreme. It seems more justifable to consider the etymologically related possessive/addressive, dative-2, ablative and comitative as part of a (perhaps secondary) grammatical series, leaving two or three (depending on the status of *-jar*) cases as postpositional or adverbial. With regard to the distinction local/grammatical, while it is clear that the possessive/addresseive and dative-2 have non-local meanings/uses, it is equally clear that some of the so-called local cases have important non-local functions. If what sets the local cases apart is the formal criterion of their use with the secondary suffixes that are also added to postpositions, then *-x′*, *-x′x′un* does look like a "defective" series. The suffix *-šša* could even fit here, but not *-ššal*. If, on the other hand, the semantics are taken into account, then these last are grammatical.

It would appear that the potential inventory of these Lak local cases is restricted in three ways: one does not occur at all *(-č̣aa)*, some appear to be facultative variants of others *(-v ~ -vu*, *-a ~ -atu*, possibly *-x ~ -xč̣in)*, and some are or are said to be no longer productive (some or all of the ad-, apud- and/or directive cases *[-c̣*, *-č̣a*, and *-n!aj* and the sublative *-lun]*, depending on the author). Apparently, as Burč̣ulaje states (1986:82), the system of local cases in Lak is undergoing change. Nonetheless, I would say that taking the formal criterion of oblique stem formation as primary, keeping in mind arguments against postpositional treatment of the local series, Lak can be said to have 10 grammatical and adverbial cases (including *-jar)* and, rejecting the differentiation of *-v* and *-vu* and *-a* and *-atu* but reserving judgment on *-x* and *-xč̣in*, 30 or 36 local cases for a total of 40 or 46.

REFERENCES

Bouda, Karl. 1949. Lakkische Studien. Heidelberg: Carl Winter.
Burč̣ulaje, G. T. 1986. Brunvata šedgenilobisa da arsebit saxelta brunebis procesebi jiritadi saḳitxebi laḳuri enaši. Tbilisi: Mecniereba.

Dirr, Adolf. 1928. Einführung in das Studium der kaukasischen Sprachen. Leipzig: Verlag der Asia minor.

Gajdarova, F. A. 1970. Obrazovanie padežej v lakskom jazyke. Lingvističeskie issledovanija, ed. by S. D. Kacnel'son, pp. 308-21. Leningrad: AN SSSR.

_____. 1977. "Funkcii mestnyx padežej v lakskom jazyke." Voprosy sintaksičeskogo stroja iberijsko-kavkazskix jazykov, pp. 273-80. Nal'čik: Institut isotrii, filologii i èkonomiki KB SSSR.

Ḫamzatov, Rasul. 1972. Ttul dağusttan. Trans. from Avar by M. Q̇. A'liev. Maxačkala: Dagestanaskoe knižnoe izdatel'stvo.

Hjelmslev, Louis. 1935. La catégorie des cas: Première partie. Universitetsforlaget. Aarhus.

Kibrik, A. E. 1978. Materialy k tipologii èrgativnosti: 2. Lakskij jazyk. (Problemnaja gruppa po èksperimental'noj i prikladnoj lingvistike, Predvaritel'nye publikacii 12.7). Moscow: AN SSSR, Institut russkogo jazyka.

Mel'čuk, Igor. 1986. Toward a Definition of Case. Case in Slavic, ed. by Richard D. Brecht and James S. Levine, pp. 35-85. Columbus: Slavica.

Murkelinskij, G. B. 1971. Grammatika lakskogo jazyka. Maxačkala: Dagučpedgiz.

_____. 1983. O poslelogax-narečijax, vyražajuščix prostranstvennye otnošenija v lakskom jazyke. Sistema preverbov i poslelogov v iberijsko-kavkazskix jazykax, ed. by N. T. Tabulova and R. X. Temirova, pp. 183-87. Čerkessk: Karačaevo-Čerkesskij naučno-issledovatel'skij institut istorii, filologii, i èkonomiki.

_____. 1989. Lakku mazral orfografialul slovar'. Maxačkala: Dagučpedgiz.

Silverstein, Michael. 1976. Hierarchy of Features and Ergativity. Grammatical Categories in Australian Languages, ed. by R. M. Dixon, pp. 112-71. [Atlantic Highlands,] New Jersey: Humanities Press.

Topuria, V. 1940. Brunebis sisṭemisatvis laḳurs enaši. Sakartvelos saxelmçipo muzeumis moambe 10-B.

Uslar, P. K. 1890. Lakskij jazyk. (Ètnografija Kavkaza:. Jazykoznanie. 4). Tbilisi: Upravlenie Kavkazskogo učebnogo okruga.

Xajdakov, S. M. 1966. Očerk po lakskoj dialektologii. Moscow: Nauka.

_____ and L. I. Žirkov . 1962. Lakku mazral wa o'rus mazral slovar'. Moscow: AN SSSR.

Xalilov, X. M., ed. 1976. Satira i jumor narodov Dagestana. Maxačkala: AN SSSR, Dagestanskij filial.

Žirkov, L. I. 1955. Lakskij jazyk. Moscow: AN SSSR.

NOTES

[1] On the other hand it can be argued that since *na* and *tanal* have the same agreement pattern in (3), they must be treated as realizing the same case, which would have to be the ergative. In Mel'čuk's (1986) sense, the ergative is a case 2 whose case 3 realizations are identical with either the nominative or genitive.

[2] The markers are the same as those given for the dative emphatic/reflexive pronouns in Table Three above. In simple sentences, the noun will agree with the nominative to which it refers, i.e., the subject of an intransitive or (facultatively) analytic transitive and the object of a transitive (obligatory if synthetic) as in the following examples based on Kibrik (1978:10-11):

8a. ppu qqatluvun-ø-aj naj ø-ur
'father goes into the house'
8b. ninu qqatluvun-n-aj naj d-ur
'mother goes into the house'

9a. nittil qqatluvun-m-aj ččąt[3] la-v-sun naj b-ur
9b. *nittil qatluvun-n-aj ččąt la-v-sun b-ur/d-ur
'mother brings bread into the house'

10a. ppu qqatluvun-m-aj ččąt la-v-sun naj ø-ur
10b. ppu qqatluvun-ø-aj ččąt la-v-sun naj ø-ur
'father brings bread into the house'

11a. ninu qqatluvun-m-aj ččąt la-v-sun naj d-ur
11b. ninu qqatluvun-n-aj ččąt la-v-sun naj d-ur
'mother brings bread into the house'

In 10/11a the focus is on the bread, in 10/11b on the parent.

[3] If the oblique stem ends in *-u*, this will front and unround to *-i* before the suffixes of this series. The *-j-* will disappear except in final position (Žirkov 1955:40). This is characteristic of the dialect Kumux (the basis of literary Lak) as opposed to that of Vicxli (the basis of earlier descriptions) (Burç̌uladze 1986:80).

[4] An entirely separate issue is the fact that certain nouns can form the inessive (and other local cases) on a base other than the oblique as derived from the genitive. Thus, for example, the genitive of *Çada* is *Çadallal* but the inessive is *Çadav* in (17) and (20). Consider also the following:

ppal 'stable'	gen. ppalnil	but: ppalav baqḫin 'to drive into the stable'
rąt 'cliff'	gen. rątnil	but: inessive rątuv (Murqqilinskij 1989) also: rątux han 'to fall off a cliff, to perish'

Uslar (1890:38) gives a number of similar examples. It is difficult to tell at this point if this is a manifestation of a tendency to shift the local cases to a simpler stem or is a lexical/adverbial phenomenon limited to toponyms and similar nouns. All the examples I have seen that are not toponymic are either buildings or topographical features. Consider also the phenomenon of the occasional loss of the genitive *-l* in compounds, e.g., *lasnaussu* 'brother-in-law' < *lasnal ussu* 'husband's brother'

[5] There aslo exists *ča* 'whence.'

[6] The post-1950 authors agree that the apudprolative *(-č̣ax)* is virtually obsolete. Burç̌uladze (1986:82) states that it only occurs with a few nouns, e.g., *qqatta* 'house' but not *č̣ira* 'wall' (cf. also Xajdakov and Žirkov 1962:403; usually the postposition *č̣arax* is used), and Murkelinskij (1971:133) omits it from his inventory and description of cases based on *-č̣a*. These authors also state that the adessive *(-c̣)* is archaic or obsolete in literary Lak (Žirkov (1955:42), Murkelinskij (1971:126), Gajdarova (1977:280), and Burç̌uladze (1986:82). Gajdarova specifies that it only occurs with *č̣ira* 'wall,' *nex* 'river,' *bak̦* 'head,' *ččan* 'leg,' *dak̦* 'heart,' *nik* 'knee,' and *xxullu* 'road,' atlthough, e.g., in the Vicxi dialect it also occurs with other nouns, e.g., *qqatta* 'house,' *murx*' 'tree' (Uslar 1890:44). Žirkov (1955:42) states: "In Kumux spoken usage the cases of this (the apud-, *-c̣*) series are little used, but they occur not infrequently in texts." Gajdarova (1977:280) aslo lists the sublative (*-lun*), *apud*-prolative (*-č̣ux)*, and all the directive (*-n!aj)* cases as unproductive. Burç̌uladze (1986:82) includes *-č̣a* with *-c̣* as being rarely heard today, their work being done by adverbs.

[7] Murkelinski (1971:125) observes that *č̣arav* or *c̣arav* (Vixli, Aštikuli dialects) corresponds to both the apud- and ad- cases.

[8] Gajdarova's (1970) inventory corresponds to Murkelinskij's but with no mention of the vocative. Her terminology follows Žirkov.

[9] 39. Uslar (1890:39) does have the following example: *Arsnaj maa'ṭra*. 'O son, do not weep!,' where the stem for the vocative is oblique. This may be an example of stem restructuring, e.g.; *ppu/buttal* 'father nom., gen.' becomes *butta/buttal* in many dialects.

[10] Uslar (1890:212) analyzes this suffix as superessive + adverbial *-nu*, e.g., *ḫaračatrajnu* 'by dint of effort' *ḫaračat* 'effort,' *ḫaračatral* = gen., *ḫaračatraj* = superessive, + *nu* = 'adv'l.

THE PARTICLE -*A* IN UDI[1]

Alice C. Harris
Vanderbilt University

Grammars of Udi describe a question particle *-a*. J̌eiranišvili (1971:121-122) notes that the question particle, *-a*, can occur with (i.e., as clitic to) nouns, adjectives, substantivized modifiers, interrogative pronouns, third person personal pronouns, adverbs, verbs, and a variety of deverbal substantives. Pančviʒe (1974:246) states that this particle occurs with all parts of speech, and Schulze (1982:136) likewise refers to *-a* as the question particle.

The present paper challenges the view that *-a* is a question particle. In section 1 the sentence types in which *-a* occurs are described, stating three conditions that must be satisfied if *-a* is to occur in a sentence. It is shown in section 2 that *-a* cannot turn a statement into a question, and it is argued on this basis that it is not a canonical question marker. Section 3 provides a partial description of the position of agreement markers in terms of three rules, and it is shown that *-a* occurs in the same positions, obeying the same three rules. As shown in section 4, *-a* also occupies the same position as agreement in copular clauses, though there are more exceptions in clauses of this type. On the basis of the facts presented in sections 3 and 4, it is argued that *-a* is best viewed as a special agreement marker.

1. Conditions on the Occurrence of *-a*. Various restrictions on the use of the question particle have been noted. J̌eiranišvili (1971:122) observes that the question particle is used only when the subject is third person singular. The question particle does not occur when the subject is first or second person or when the subject is plural. Although J̌eiranišvili provides no examples to illustrate this, elicited examples confirm his statement:

(1) (a) okt'omber-a evaxt'-un taysa?[2]
Okt'omber-Dat when-2.Sg go
'When are you going to Okt'omber [village]?'
(b) okt'omber-a evaxt'-a̲ taysa?
Okt'omber-Dat when-Q go
'When is s/he going to Okt'omber [village]?'

(c) okt'omber-a evaxt'-yan taysa?
Okt'omber-Dat when-1.Pl go
'When are we going to Okt'omber [village]?'
(d) okt'omber-a evaxt'-nan taysa?
Okt'omber-Dat when-2.Pl go
'When are you going to Okt'omber [village]?'
(e) okt'omber-a evaxt'-q'un taysa?
Okt'omber-Dat when-3.Pl go
'When are they going to Okt'omber [village]?'

(2) (a) zu ma-z aš-besa?
I where-1.Sg work-do
'Where do I work?'
(b) ma-n aš-besa?
where-2.Sg work-do
'Where do you work?'
(c) me xinären ma-a aš-besa?
this girl where-Q work-do
'Where does this girl work?'
(d) ma-yan aš-besa?
where-1.Pl work-do
'Where do we work?'
(e) ma-nan aš-besa?
where-2.Pl work-do
'Where do you (Pl) work?'
(f) vi xunči-n-γo-n ma-q'un aš-besa?
your sister-Ext-Pl-Erg where-3.Pl work-do
'Where do your sisters work?'
(3) (a) xinär ma-a taysa?
girl where-Q go
'Where is the girl going?'
(b) xinär-m-ux ma-qun taysa?
girl-Ext-Pl where-3.Pl go
'Where are the girls going?'

Pančviʒe (1974:246), comparing Udi -*a* to the Old Georgian question particle of the same form, notes that the Udi particle has the peculiarity that it occurs **even** when there is a question word (interrogative pronoun, adjective, adverb) present. The examples cited above confirm that the particle does occur when a question word is present. What is perhaps more surprising is that the particle occurs **only** when a question word is present. In simple yes/no questions, the question particle does not occur, as shown in the following examples.

(4) (a) vi baba ar-e-ne.
your(Sg) father come-Aor-3.Sg
'Your father came.'
(b) vi baba ar-e-ne?
your(Sg) father come-Aor-3.Sg
'Did your father come?'
(5) (a) aššlax b-e-ne.
matter do-Aor-3.Sg
'S/he took care of the matter.'
(b) aššlax b-e-ne?
matter do-Aor-3.Sg
'Did s/he take care of the matter?'

There is no difference between the statements, (a), and the questions, (b), except intonation. Comparison of the (a) and the (b) examples shows that the particle -*a* does not occur in yes/no questions; examination of text examples confirms this. The examples in (6) show that the particle cannot be added optionally to a yes/no question.

(6) (a) xorag k-e-nu?
food eat-Aor-2.Sg
'Did you eat the food?'
(b) *xorag k-e-nu-a?
food eat-Aor-2.Sg-Q
(c)*xorag k-e-a?
food eat-Aor-Q
(d) *xorag-a k-e-nu?[3]
food-Q eat-Aor-2.Sg

The particle *-a* is not found in yes/no questions; it is not found without a question word. Put differently, the so-called question particle cannot itself form a question.

We should point out that the question particle *-a* does not occur even when all of the above-stated conditions are met, unless the clause is a question. All other necessary conditions may be met, yet *-a* not occur, in the sort of relative clause that is formed with a relative pronoun, which in Udi may be identical to a question word.[4] (7) illustrates a relative clause formed with a relative pronoun, identical to the question pronoun; note that no question particle occurs in either clause.

(7) šin zavod-a aš-ne-bsa, š̄um-enk' mia-ne are.
who/Erg factory-Dat work-3.Sg-do bread-for here-3.Sg come
'Those who work in the factory come here for bread.'

Thus, it is not the question word itself that conditions the occurrence of the particle *-a*, but rather the fact that the clause in which it occurs expresses a content question.[5]

Pančviʒe has stated the condition on the occurrence of the question particle in the following way (1974:149): the question particle, rather than the agreement marker, occurs when the subject is third person and there is a question word in the sentence. This statement is incomplete in several ways. First, as stated by J̌eiranišvili (1971:122) and as shown above, it is only with third person **singular** subjects that *-a* occurs. Second, question words condition the particle only in a content question, not in relative clauses. Third, inversion verbs are a systematic exception; that is, with an inversion verb, agreement, not *-a*, must occur, as shown in the sentences below.

(8) met'in ek'a-t'u aba?
s/he/Erg what/Absl-3.Sg know
'What does s/he know?'

(9) šux t'u ak'sa me xinär-en?
who/Dat 3.Sg see this girl-Erg

'Whom does this girl see?'

(10) šin t'u aba me äzbär-a?
who/Erg 3.Sg know this poem-Dat
'Who knows this poem?'

These examples show that even when the subject is third person singular, if the verb is an inversion verb, the third person singular agreement marker, *t'u*, occurs, rather than the question particle *-a*.[6] In Udi there is a distinct set of agreement markers for use in inversion verbs, as set out in the table below (based on J̌eiranišvili 1971:79-80, Pančviʒe 1974:133, 138, Schulze 1982:170, 174; some variants not included here):

	Direct verbs	Indirect verbs
1. Sg	-zu, -z	-za
2. Sg	-nu, -n	-va
3. Sg	-ne	-t'u
1. Pl	-yan	-ya
2. Pl	-nan	-vạ
3. Pl	-q'un	-q'o

As examples (8-10) show, inversion verbs in Udi are chiefly characterized by the use of this set of distinctive agreement markers, rather than by distinctive syntax or case marking patterns, as in other languages of the Caucasus.[7] The position of agreement markers relative to other constituents is discussed below.

A fourth emendation that must be made to Pančviʒe's generalization is that certain tense forms are an optional exception; that is, with these tense verb forms, agreement, rather than *-a*, may optionally occur.

(11) (a) ek'alu eγ-al-le vi viči?
why come-Fut-3.Sg your brother/Absl
'Why will your brother come?'

(b) ek'alu eγ-al-a vi viči?
why come-Fut-Q your brother/Absl
'Why will your brother come?'
(c) ek'alu-a are vi viči?
why-Q come your brother/Absl
'Why did your brother come.'

Examples (11a) and (b) are in the future tense that is based on a participle. In (11a) the final suffix is an agreement marker, *-ne*, assimilated to the preceding *l;* in (11b) the agreement marker does not occur, and in its place is the question particle. Thus the question particle is optional in this form. Example (c) shows the question particle, which is required in most tenses, here illustrated with the aorist.

We can summarize the conditions on the occurrence of *-a* as follows: The question particle *-a* occurs if and only if:

1. The clause is a question containing a question word.
2. The subject of the sentence is third person singular.
3. The verb of the clause is a direct verb, not an inversion verb.

Occurrence is optional in certain tenses. Note that these conditions state only **in which sentences *-a* will occur**; they do not tell us where in the sentence it will appear. The latter is discussed below in section 3.

2. Is *-a* a Question Particle? The view that *-a* is a question particle leaves unexplained the curious fact that when it occurs, the third person subject marker *-ne* fails to occur (Ǯeiranišvili 1971:122, Pančviʒe 1974:149).

Since the so-called question particle *-a* occurs only in content questions, it undoubtedly contributes to indicating that the clause in which it appears expresses a question. We can easily see that it was on the basis of this limitation to questions that the particle was originally designated a question particle.

By the same token, since the particle occurs only with third person singular subjects, it indicates that the clause in which it appears has a third person singular subject. That is, the particle fulfills, among others, the functions of an agreement marker.

The term "question particle" is ordinarily used for a particle that can turn a statement into a question. The sentences in (6) show that *-a* cannot do this in Udi. It cannot occur at all in yes/no questions. It is not the question particle that makes content questions into questions, as shown by the fact that content questions (8-10), in which it does not occur, are nevertheless questions. Since *-a* cannot simply be added to any statement to turn it into a question, it does not fulfill the functions ordinarily associated with a question word. From a synchronic point of view, we might, therefore, consider it to be a special agreement marker, reserved for certain types of question sentences. This view is supported by a further fact, discussed in section 3. (I continue to refer to *-a* as the question particle and to gloss it as Q.)

3. The Position of the Agreement Marker in Udi. In this section it is shown that, although the agreement marker in Udi occurs in a wide variety of positions under different circumstances, the so-called question particle occurs in the position that would otherwise be occupied by the agreement marker. The description is limited to the Vartašen dialect (including Okt'omber), as the dialects differ in this regard.

The position of the agreement marker is a complicated matter that still cannot be fully described. J̌eiranišvili (1971:83-84) states that (1) with stative verbs the agreement marker is always final, as in *arci-ne* 's/he sits,' *arci-te-ne* 's/he does not sit,' and numerous other examples, and (2) with transitive and intransitive dynamic verbs the agreement marker is in various positions. He observes that the agreement marker is never the initial element; if it is a prefix, it always follows some other prefix (e.g., *te-ne-esa* 's/he does not come'). In composite verbs the agreement marker is either between the two bases (e.g., *aš-ne-besa* 's/he works') or after both (e.g., *aš-besa-ne* 'having worked'). Clearly this description leaves many questions about the precise position of the marker unanswered.

Pančviʒe (1974:148-150) observes that the agreement marker is ordinarily what he calls an infix; that is, it is both preceded and followed by other parts of the verb complex. However, with certain verb forms it may be the [final] suffix. His paradigm of verb forms

(Pančviʒe 1974:155) helps to make this statement more specific, for the forms listed there show the agreement marker as the final or penultimate suffix only in the subjunctive I, subjunctive II, future II, and conditional (e.g., *kar-x-al-zu* 'I will live,' where *-zu* marks the first person singular subject). He points out that the agreement marker often precedes the verbal base, as in *te-z-cam-p-i* 'I did not write,' literally 'not-1.Sg-writing-said,' where the verbal base is *cam-p-* 'write, writing-say.' He adds that sometimes the agreement marker is entirely outside the verb, as in (12).

(12) sa q'oǰa kaft'ar-re pasč'aγun-t'ọγọl esa
one old woman-3.Sg king come
'An old woman comes to the king.'
(Bežanov 1988:8,7; cited by Pančviʒe 1974:149)

Pančviʒe (1974:149) suggests that we would have expected instead *sa q'oǰa kaft'ar pasč'aγun-t'ọγọl e-ne-sa*, with the agreement marker between the two elements of the verb base, as it ordinarily is.

Schulze (1982:168-169) observes that the agreement markers (PZ) may occur in any one of the following locations, where V represents the verb stem, T a tense marker, Vn a nominal extension of the verb, and HV an auxiliary:

(1) V-T-PZ
(2) V-T^1-PZ-T^2
(3) V^1-PZ-V^2-T
(4) Vn-PZ-(H)V-T
(5) Vn-(H)V-T-PZ
(6) ...X-PZ-V-T

Some rules of thumb are also provided, such as the fact that the suffixes *-s-a*, which together mark the present tense, cannot be split by the agreement marker. Schulze further suggests that his position (5) is limited to special contexts, such as the hortative/imperative.

While all three authors have contributed in important respects to our understanding of the position of agreement markers, it remains impossible to predict where the marker will occur in any given clause. While we are aware of the **possibilities** that exist, we still cannot say exactly where the marker **will** occur under specific circumstances.

Below are three rules that predict where the agreement marker will go. These are not exhaustive, they do not cover all instances; but they are correct for all clauses that meet these conditions. Thus, they are only a first step toward a positive statement of the rules:

1. Agreement is final in the verb if it is in the future II (participial future) form.[8]
2. Agreement is enclitic to the negative particle.
3. Agreement is enclitic to the questioned constituent.

Examples in (13) illustrate the application of Rule 1, (14) that of Rule 2, and (15) that of Rule 3, each in environments with no rule conflict.

(13) (a) bez vič-en ašb-al-le zavod-a.
my brother-Erg work-Fut-3.Sg factory-Dat
'My brother will work in the factory.'
(b) xorag box-al-zu.
food/Absl prepare-Fut-1.Sg
'I will prepare food.'

(14) (a) mählina xod te-ne bošt'esa.
yard/in tree Neg-3.Sg plant
'S/he is not planting the tree in the yard.'
(b) šet'in te-t'u aba ạil-ax etär kalab-al-a.
he/Erg Neg-3.Sg know child-Dat how raise-Fut-Q
'S/he does not know how to raise a child.' '...how she will raise a child.'[9]

(15) (a) ek'a-n besa?
what/Absl-2.Sg. do
'What are you doing?'
(b) tivliz-a šux-va čalxesa?
Tbilisi-Dat who/Dat-2.Sg know
'Whom do you know in Tbilisi?'
(c) tivliz-a šin-t'u čalxesa?
Tbilisi-Dat who/Dat-3.Sg know
'Who in Tbilisi knows you?'
(d) mano muz-nu äytexa.
which language-2.Sg speak
'Which language do you speak?'

These examples show that the rules stated above apply to both sets of agreement markers ((14b) and (15b, c) are inversion verbs).[10] (15d) establishes that it is the questioned constituent ('which language') to which *-a* cliticizes, not necessarily the question word itself. In the examples above, placing the agreement marker in positions other than those allowed by the rules produces ungrammatical sentences.[11] Interaction of these rules is discussed below.

The question particle *-a* occurs in exactly the same positions as the agreement markers. (16) shows *-a* following the participial future, parallel to (13). (17) parallels (14), illustrating the question particle following the negative particle, *te-*. (18) shows the question particle in the same position that the agreement marker occupies in (15), following a question word.

(16) (a) šin k'alk'-al-a lašk'on-a?
who/Erg invite-Fut-Q wedding-Dat
'Who will invite [people] to the wedding?'
(b) šux k'alk'-al-a lašk'on-a?
who/Dat invite-Fut-Q wedding-Dat
'Whom will s/he invite to the wedding?'

(17) ek'alu te-a are vi viči?
why Neg-Q come your brother

'Why didn't your brother come?'

(18) (a) šu-a are mia?
who/Absl-Q come here
'Who came here?'
(b) mano xinär-a are?
which girl-Q come
'Which girl came?'

Sentences (16-18) show that the question particle occurs in the same positions as agreement markers illustrated in (13-15).[12] In both instances, these are the only positions these markers can occupy.[13] (18b) establishes that it is the questioned constituent ('which girl') to which -*a* cliticizes, not necessarily the question word itself (just as for agreement).

In some clauses, the conditions for more than one rule are met at the same time. Rule 1 takes precedence over rules 2 and 3; rule 2 takes precedence over 3.[14] That is, if the conditions on two rules are met at the same time (in the same clause), then it is the rule with the lower number that applies. This is shown first for agreement markers in the examples that follow. In (19), the conditions on Rules 1 and 2 are both met, and it is Rule 1 that applies.[15]

(19) (a) mählina xod nut' bošt'-al-le.
yard tree Neg plant-Fut-3.Sg
'S/he will not plant a tree in the yard.'
(b) *mählina xod te-ne bošt'-al.
yard tree Neg-3.Sg plant-Fut
'S/he will not plant the tree in the yard.'
(c) damnun nut' bak-al-t'u.
tomorrow Neg can-Fut-3.Sg
'Tomorrow s/he cannot [come here].'

In case the conditions on Rules 1 and 2 are both met, a special negative particle, *nut'*, is used, as in (19a); indeed, its use is conditioned by the coinciding of negation and the participial future. (19b) shows that it is ungrammatical to cliticize agreement to the usual negative particle, *te*, whether or not agreement is also attached to the participial future; this contrasts with (14), where agreement does cliticize to *te*.[16]

In (20), the conditions on Rules 1 and 3 are both met, and it is Rule 1 that applies.

(20) ek'a b-al-lu?
what/Absl do-Fut-2.Sg
'What will you do?'

This shows that when the clause is both in the participial future and contains a question word, it is the future verb form to which agreement attaches.

Rule 2 takes precedence over Rule 3; (21) illustrates the interaction of these rules in a single clause.

(21) (a) ek'alu te-n mạγexa?
why Neg-2.Sg sing
'Why don't you sing?'
(b) evaxt' te-t'u baksa mia eγes?
when Neg-3.Sg can here come
'When can s/he not come here?'

This shows that when the negative particle *te* appears in the same clause with a question word, it is the negative to which the agreement marker cliticizes. (21a) shows this with the direct construction, and (21b) with inversion.[17]

It remains to show that the interaction of Rules 1-3 is the same for the question particle as for agreement markers. In order to test the interaction of Rules 1 and 2 with the question particle, a future negative question is required. This is not a very natural form because the circumstances which the form would fit are not very natural, and speakers substituted others on my attempts to elicit this. In the other instances the results are clear. Rule 1 takes precedence over Rule 3:

(22) ma bošt'-al-a̲?
where plant-Fut-Q
'Where will s/he plant it?'

For the question particle, (23) illustrates the interaction of Rules 2 and 3.

(23) (a) ek'alu te-a̲ mąγexa?
why Neg-Q sing
'Why doesn't s/he sing?'
(b) mąγ ši te-a̲-kefilesa?
song who/Gen Neg-Q-like
'Who doesn't like singing?'
(c) šu te-a̲-are?
who/Absl Neg-Q-come
'Who did not come?'

Thus, as far as can be determined, the interaction of these three rules is the same for the question particle as for agreement markers, and it appears that the behavior of this particle is identical to that of the agreement markers.

Since the question particle *-a* cannot turn statements into questions, it seems inappropriate to consider it a question particle at all. Since it occurs only with third person singular subjects, it functions in part as a person marker. Since it is treated as a person marker by the Rules 1-3, which state the placement of such markers, it seems most appropriate to consider *-a* a marker of the third person singular subject, with the additional limitations described above.

4. The Copula. Ǯeiranišvili (1971:122) has remarked that when the particle *-a* takes the place of the agreement marker, it expresses also the meaning of the copula. Let us look at the ordinary expression of the copula in Udi. The copula is generally not realized (is realized as zero) in Udi, as in the following examples.

(24) (a) zu doxt'ur-zu.
I/Absl doctor-1.Sg.
'I am a doctor.'
(b) šono doxt'ur-re.
he/Absl doctor-3.Sg
'S/he is a doctor.'
(c) me čubux doxt'ur-re.
this woman/Absl doctor-3.Sg
'This woman is a doctor.'

(25) un p'ạq'o usen-un-nu.
you.Sg forty year-Gen-2.Sg
'You are forty years old.'

(26) t'e čubux xib äil-un nana-ne.
that woman three child-Gen mother-3.Sg
'That woman is the mother of three children.'

In each example the agreement marker, agreeing in person and number with the subject, is enclitic to the predicate nominal, whether that is a noun or other constituent. (27) shows that the predicate nominal can also be a participle.

(27) ari-zu.
arrived-1.Sg
'I have arrived.'

In yes/no questions, we find the same pattern:

(28) šono nana-ne?
she/Absl mother-3.Sg
'Is she a mother?'

(29) mono bez k'inigi-ux-ne?
it my book-Pl-3.Sg
'Are those my books?'

(30) ema usen-un-nu?
howmany year-Gen-2.Sg
'How old are you?'

In a sense, the agreement marker fulfills the function of the copula in Udi. There are, however, contexts in which 'be' is expressed as *bu-* or as *bak-*, as in (31) and (32):

(31) (a) maɣazina ek'a bu-va-q'-sa bu-ne.
store/Dat what/Absl want^1-2.Sg-want^2-Pres be-3.Sg
'Whatever you want is in the stores.'
(b) xinär bez bu mano bu-ne qo usen-un.
girl my be who be-3.Sg five year-Gen
'I have a girl who is five years old.'

(32) okt'omber-a te-n bak-e?
Okt'omber-Dat Neg-2.Sg be-Aor
'You haven't been in Okt'omber?'

We can state this generalization regarding the placement of the agreement marker in copular clauses as Rule 4:

> 4. Agreement is enclitic to the predicate nominal in a clause with a present tense copula (zero).

The so-called question particle, *-a*, behaves as an agreement marker in this respect also. That is, in a copular clause, *-a* cliticizes to the predicate nominal:

(33) eq'q'ara st'udent' xinär-a?
howmany student/Absl girl/Absl-Q
'How many students are girls?'

Since *-a* occurs only in content questions (see section 1 above), the conditions on Rule 3 are always met, and this rule generally takes precedence over Rule 4. ((33) is an exception to the generalization involving interaction of the two rules.)

(34) mia mano čubux-a doxt'ur?
here which woman/Absl-Q doctor
'Which woman here is a doctor?'

(35) šu-a doxt'ur?
who/Absl-Q doctor
'Who is a doctor?'

(36) mano-a šavat'?
which-Q pretty
'Which one is pretty?'

(37) me gazet' evaxt'-un-a?
this newspaper when-Gen-Q
'When is this newspaper from?' 'What day's newspaper is this?'

(38) nana ma-a?
mother where-Q
'Where is mother?'

(39) hava etär-a?
weather whatkind-Q
'What kind of weather is there?' 'What is the weather like?'

We must add that in some instances agreement, rather than *-a*, is used in content questions involving a copula. (40) is parallel to (30); but because (40) has a third person singular subject, we would expect *-a*, rather than *-ne*.

(40) ema usen-un-ne?
howmany year-Gen-3.Sg
'How old is he?'

Although there seem to be a few peculiarities in the use of -*a* in the copular clause, on the whole we can say that -*a* behaves like agreement in this respect, too.[18] Both agreement and -*a* obey Rule 4.

5. Conclusion. Although -*a* is restricted to questions, it does not **form** questions; it is therefore an error to consider it a question particle. The view that this is a question particle leaves unexplained the fact that agreement occurs in all sentences except those containing -*a*. This view offers no explanation of the fact that -*a* occurs only when the subject is third person singular, just as the subject markers -*ne* and -*t'u* (inversion) do.

The view suggested here is that -*a* is a variant of -*ne*, used just in the context of a question word. The view that -*a* is essentially a person marker explains both the fact that no other agreement marker occurs when -*a* is present and the restriction of -*a* to one person and number. This position is supported by the fact that -*a* occurs in the same positions as agreement markers, following Rules 1-3, and by the fact that -*a* behaves like an agreement marker in copular clauses.

REFERENCES

Bežanov, Mixail. 1988. Rust'am [text of a folk tale]. *Sbornik materialov dlja opisanija mestnostej i plemen Kavkaza* 12. Tbilisi.

Harris, Alice C. 1984. Case marking, verb agreement, and inversion in Udi. *Studies in Relational Grammar, 2*, ed. by David M. Perlmutter and Carol G. Rosen, 243-258. Chicago: University of Chicago Press.

J̌eiranišvili, Evg. 1971. *Udiuri ena* [The Udi language]. Tbilisi: Universit'et'i.

Pančviʒe, Vl. 1974. *Uduri enis gramat'ik'uli analizi* [A grammatical analysis of the Udi language]. Tbilisi: Mecniereba.

Schulze, Wolfgang. 1982. *Die Sprache der Uden in Nord-Azerbajdžan*. Wiesbaden: Harrassowitz.

Sixaruliʒe, Tariel. 1987. *Arsebiti saxelis bruneba udiurši* [Declension of nouns in Udi]. Dissertation. Tbilisi: Linguistics Institute.

NOTES

[1] This material is based upon work supported in part by the National Science Foundation under Grant No. BNS-8712111, by the International Research and Exchanges Board under the ACLS-Academy of Sciences Exchange, and by the

National Endowment for the Humanities under a Fellowship for University Teachers. This paper is based upon work with consultants from the village Okt'omber and one from Niǰ, and all examples except (12) are elicited. I find nothing in published texts to suggest that the use of the question particle in Vartašen is any different. I am grateful to my consultants Dodo Misk'ališvili, Luiza Nešumašvili, Simon K'isbabašvili, and Alvina Mat'ilian, and to Tina Misk'ališvili, who helped me in many ways.

[2] In glossing examples, I have used traditional names of cases and grammatical categories uncritically, including the tradition of calling *-a* a question particle. Pharyngealized vowels are marked with a subscript dot. The following abbreviations are used: *1.* "first person," *2.* "second person," *3.* "third person," *Erg* "ergative case," *Absl* "absolutive case," *Dat* "dative case," *Gen* "genitive case," *Sg* "singular," *Pl* "plural," *Pres* "present tense," *Fut* "participial future tense, future II," *Aor* "aorist (II) tense," *Mas* "masdar," *Ext* "root extension," *Neg* "negative particle," *Q* "question particle." Generally words are analyzed into constituent morphemes only to the extent that such analysis is relevant to the points being made. Morphemes analyzed in the Udi original are separated by hyphens and their glosses by hyphens; glosses of morphemes that are not separated in the Udi original are themselves separated by a slash.

[3] The form

xorag-ax k-e-nu?
food-Dat eat-Aor-2.Sg

is fully grammatical, as Udi permits direct objects to be either in the absolutive or in the dative case. Ordinarily *xorag-a* would be an alternate form of *xorag-ax*, and one would expect (6d) therefore to be grammatical. We may assume that the syncretism of the reduced form of the dative and the form with enclitic *-a* prevents (6d) from being grammatical. In any case, my informant would not accept it.

[4] The relativizing particle *-te* may be added to the question word in headless relative clauses but usually is not; the relativizing particle is usually added to the question word in a headed relative clause but may be omitted.

[5] The question particle does occur in embedded questions; an example is the second clause in (14b) below.

[6] Many or most verbs that govern inversion morphology in the Vartašen (and Okt'omber) dialect do not in the Niǰ dialect (Pančviʒe 1974:142). This accounts for the fact that a verb that takes an agreement marker in a question in Okt'omber

šin t'u ibaksa bez äytpesun-a?
who/Erg 3.Sg understand my speaking-Dat
'Who understands my talk?'

may permit a question particle instead in Niǰ:

šin-a̲ ibaksa bez exlə̣t-ə̣?
who/Erg-Q understand my speaking-Dat
'Who understands my talk?'

[7] For most inversion verbs, these consultants consistently put experiencers in the ergative case and stimuli in the absolutive or dative case; this is one of only two patterns found in texts (see Harris 1984). A few verbs, such as 'have' and 'like' govern a different pattern with experiencers (possessors) in the genitive.

[8] There are additional tense forms to which the agreement marker seems to be consistently enclitic, but I have investigated in detail the interaction of only the future with other rules.

[9] Example (14b) uses a question word, *etär* 'how', and a question particle because it contains an indirect question, 'how she will raise a child'.

[10] The following sentences from Nij̆ illustrate the same principle at work there:

Rule 1 okt'ombеr-in-a evax taγ-al-n̲u̲
Okt'ombеr-Ext-Dat when go-Fut-2.Sg
'When will you go to Okt'omber?'

Rule 2 ukes te-z̲ baksa.
eat/Mas Neg-1.Sg can
'I cannot eat.'

Rule 3 k'ä-n̲ bsa?
what/Absl-2.Sg do
'What are you doing?'

[11] The position of the agreement marker relative to the verb is free (except where Rule 1 applies). The sentences below show that the verb may or may not immediately follow the agreement marker:

finax ma-n̲ bųγabe?
wine/Dat where-2.Sg get
'Where did you get wine?'

ma-n̲ finax bųγabe?

Nevertheless, it is usual for the verb to follow or contain the agreement marker.

[12] The sentences below show that the same rules apply in the Nij̆ dialect:

Rule 1 vi viči het'aynak eγ-al-a?
your.Sg brother why come-Fut-Q
'Why will your brother come?'

Rule 2 hevädä eyes te-ne bako mia?
when come Neg-3.Sg can here
'When can he not come here?'

Rule 3 het'aynak-a hare vi viči?
why-Q come your.Sg brother
'Why did your brother come?'

[13] The position of the question particle relative to the verb is free (except where Rule 1 applies). The sentences below show that the verb may or may not immediately follow *-a*.

išen šux-a pešk'äšbe tängina?
man/Erg who/Dat-Q give money/Dat
'To whom did the man give money?'

šux-a išen pešk'äšbe tängina?

šux-a tängina pešk'äšbe išen?

[14] Rule 2 also takes precedence over Ǯeiranišvili's rule, cited above, that agreement is final in stative verbs; compare the sentence below:

nana arci-ne
mother/Absl sit-3.Sg
'Mother is sitting.'

nana te-ne arci.
mother/Absl Neg-3.Sg sit
'Mother is not sitting.'

[15] There are exceptions to the generalization that Rule 1 has precedence over Rule 2. My consultants offered sentence (i) below, and Sixaruliʒe 1987:32 gives (ii), as well as another example, each part of a proverb.

(i) me xe-n-ax ųγes bakal te-ne.
this water-Ext-Dat drink/Mas can/Part Neg-3.Sg
'This water is not drinkable.'
(ii) ...uq'eyn-ax xurubal te-ne.
bone-Dat Neg-3.Sg
'...he does not make the bones dusty.'

There may be a meaning difference involved. When presented with (iii) and (iv), one consultant suggested that the first is a direct statement, the second a reply to something else.

(iii) bez k'ua nut' taš-al-zu vax.
my house/Dat Neg take-Fut-1.Sg you.Sg/Dat
'I will not take you in my house.'
(iv) bez k'ua taš-al te-zu vax.
my house/Dat take-Fut Neg-1.Sg you.Sg/Dat

[16] In Niǰ, on the other hand, it appears that Rule 2 may take precedence over Rule 1:

šin dava-ne-bsa, k'wä te-z bark'-al.
who/Erg fight-3.Sg-do house/Dat Neg-1.Sg allow-Fut
'Whoever fights I will not allow into the house.'

Compare this with the same sentence from Okt'omber:

šin-te dava-ne-saksa, k'wa nut' bark'-al-yan.
who/Erg-Rel fight-3.Sg-do house/Dat Neg allow-Fut-1.Pl
'Whoever fights we will not allow into the house.'

[17] The same generalization holds for a third set of agreement markers, used with only a handful of verbs, including 'have' and 'like':

evaxt' te-vi bu vädä?
when Neg-2.Sg have time
'When do you not have free time?'

[18] A special instance of the use of -*a* occurs in either/or questions; note that this type of question is distinct from both the yes/no question and the content question.

In either/or questions not involving the copula, the word *yoxsam* 'or' is used, as in (i) and (ii).

(i) un eγal-lu yoxsam mandal-lu?
you.Sg go-2.Sg or stay-2.SG
'Will you go or will you stay?'
(ii) vi baba are-ne yoxsam te?
your.Sg father come-3.Sg or Neg
'Did your father come or not?'

This may also be used in a clause that involves a copula:

(iii) täzä-ne yoxsam bisi?
new-3.Sg or old
'Is it new or old?'

A different alternative uses *-a*, rather than *yoxsam*. Note that this is the only use of *-a* we have seen that does not involve a question word in the same clause.

(iv) me k'oǰ k'ic'k'e-a kala?
this house small-Q large
'Is this house small or large?'

CRIMEAN TATAR AND KRYMCHAK: CLASSIFICATION AND DESCRIPTION

Maria S.Polinsky,
University of Southern California, Los Angeles

ABSTRACT

The conventional classification of Turkic languages of the Crimea is analyzed, and an alternative classification is suggested which distinguishes between the two variants of Crimean Tatar (Orta and Coastal Crimean Tatar) on the one hand, and Steppe Crimean Tatar on the other. While the former variants are related to the Oghuz grouping of Turkic, the latter belongs to the Kypchak grouping. Thus, the homogeneity of the accepted Crimean Tatar language proves problematic. It is argued that, in the case of closely related languages, the classification should rely on both lexical correspondences and correspondences between relational morphemes. The latter criterion is credible for closely related languages, as the time of the possible change is quite short. In the Crimean Tatar case, the lexicostatistical and morphological data prove consistent. A subdialectal variant of Orta/Coastal Crimean Tatar, formerly spoken by the Sephardic Jewish group of the Crimea, is described. This variant, known as Krymchak, differs from Crimean Tatar proper in that it has a greater mixture of the Oghuz and Kypchak elements, as well as in the greater variation across local variants and/or idiosyncrasies of individual speakers. This can be due to sociolinguistic factors. To describe the variation between Crimean Tatar and Krymchak, the notion of ethnolect is proposed. Ethnolect is defined as the variant of the single or predominant language as it is spoken by a distinct and non-assimilated minor group within the numerically prevalent population.

This paper is motivated by two related concerns. First, an attempt is made to classify the Turkic vernaculars of the Crimean Peninsula. Second, a brief description of Krymchak, an ethnic version of Crimean Tatar which has been little studied before, is given[1]. This description is also accompanied by a discussion of the theoretical notion of ethnolect.

Section I discusses the problems of genetic classification of Turkic languages or vernaculars of the Crimea. Section II is a general description of Krymchak; in this description, comparison to Crimean Tatar proper is also made. Section III discusses the interaction of languages and language versions, namely, (a) the structural relationship between Crimean Tatar proper and Krymchak, (b) the interaction of Oghuz and Kypchak language elements in the Turkic languages of the Crimea. Section IV presents the notion of ethnolect.

I. CLASSIFICATION

1. Introduction. Traditional classifications of Turkic languages distinguish the following groups within the Turkic subfamily of Altaic (Deny 1959; Voegelin and Voegelin 1977:339):

(i) Chuvash
(ii) Kypchak (Central and Western Turkic)
(iii) Karluk (Eastern and Northern Turkic)
(iv) Oghuz (Southern/Southwestern Turkic).

There have been two basic approaches to the languages of the Crimea. According to the conventional view followed in Soviet Turcology after È.V.Sevortjan (1966:234-235), the main bulk of the Turkic population in the Crimea used to be speakers of Crimean Tatar, which is a Kypchak language. This view distinguishes three dialects within Crimean Tatar, roughly located in horizontal belts from the north to the south of the peninsula: Steppe, Middle (Orta Julaj) and Coastal. The classification is also followed in Hesche and Scheinhardt 1974. The other view (Doerfer 1959:369-373) distinguishes three languages: Crimean Nogay, Crimean Tatar proper and Central Crimean Turkish. The latter language is Oghuz, while the two others are considered Kypchak. In Doerfer's classification, Crimean Nogay roughly corresponds to the Steppe dialect, while each of the two other languages overlaps with both of the remaining dialects.

These classifications need to be tested against the empirical data collected from fieldwork with Crimean Tatar speakers. The study of the Crimean Tatar situation has certainly been hindered by the political situation of the ethnic group, which was deported from the Crimea during World War II at Stalin's order. The deportation triggered language attrition and dialect change. The relative revival of the language in the 1960s, represented by the publication of two periodicals (the magazine *Yildiz* [Star] and the newspaper *Lenin Bayraghi* [Lenin's Banner]), is associated with a different language situation, in which former dialect distinctions are nearing extinction. Despite the loss of dialectal divisions, ethnic awareness is retained, and these ethnic distinctions correlate with the former dialectal division.

The following problems arise regarding any classification within the Turkic group:

(1) a. the languages are so genetically close that any classification risks inadequacy;

b. as a result, separate languages are distinguished despite the fact that the conventional conception of language does not seem to apply;

c. in model genetic classifications, not enough emphasis is placed on grammatical elements, especially relational morphemes which allow languages to be differentiated.

2. Comparing basic word lists. To verify the distinctions between the tentative languages, basic word-lists were collected from speakers who identified themselves as speakers of a certain dialect. Table 1 shows the results of a test where 100-word Swadesh lists were compared for the three tentative dialects of Crimean Tatar, and also for Ottoman Turkish and Gagauz (the Oghuz group), Nogay and Kumyk (the Kypchak group). (Basic vocabulary lists for the three Crimean Tatar dialects, obtained from native speakers, are presented in Appendix I.)

The results indicate relationships which differ from the ones summarized above. First, the differences in the basic vocabulary are generally quite subtle, which suggests that what is conventionally treated as a set of languages is rather a set of dialect clusters. The reason for treating these as languages is primarily non-linguistic and has to do with the ethnic identity of the respective groups.

TABLE 1: Percentages of basic vocabulary in pairs of Turkic languages calculated for 100-word Swadesh list.

		CrTC	CrTS	Tur	Gag	Nog	Kum
CrTO	91%	82%	84%	83%	80%	80%	
CrTC		77%	90%	89%	77%	76%	
CrTS				74%	78%	87%	87%
Tur					89%	78%	77%
Gag						77%	75%
Nog							88%

In addition to the close genetic relationship, much linguistic contact can be speculated. The contact was definitely facilitated by the closeness of the original relationship. To exclude the effect of contact interference in the data, the following technique was used. Given a pair of closely related languages (A and B), their relationships to other languages are checked. If one of the languages in the pair shows an unusually high percentage of cognates with some other language (C), while the other does not, this former percentage is attributed to language contact which followed the split. Thus, if A and C show a higher percentage of cognates than B and C, the relationship of A and B together to C is assessed by the lower percentage, i.e., the one in the B-C pair.

In Table 1, this can be illustrated by the pair Crimean Tatar Orta — Crimean Tatar Coastal (91% cognates). Once these two languages are defined as a cluster, their relationships with other languages are estimated on the basis of the lower of the two cognate percentages obtained. Accordingly, with Turkish, it is 84%, rather than 90%; the latter is explained by later language contacts. By the same token, the cluster has 83% cognates with Gagauz, 77% cognates with Nogay and 76% with Kumyk.

This analysis reveals the following clusters:

(2) a. Crimean Tatar Orta — Crimean Tatar Coastal
    b. Turkish — Gagauz
    c. Nogai — Kumyk
    d. Crimean Tatar Orta/Coastal — Turkish/Gagauz
    e. Crimean Tatar Steppe — Nogai/Kumyk

Of these, (2b) and (2c) are consistent with the existing classification and represent respectively the Oghuz and Kypchak group. Crimean Tatar, on the other hand, does not conform to the notion of a homogeneous language. Rather, it is a union of dialects apparently related to two different groups: Oghuz (2d) and Kypchak (2e). In principle, the resulting classification resembles the one proposed by Doerfer. However, contrary to Doerfer's view, Crimean Tatar Orta is not recognized as a separate language, since it is closely related to the Coastal dialect.

The linguistic clusters thus obtained reflect the history of the Turkic speaking groups in the Crimea. There seem to have been two migration directions, namely the Kypchak wave from the north-east and the Oghuz wave from the south-east, via Iran and the Caucasus. The resulting classification is summarized in Figure 1. The dotted line indicates later language contacts.

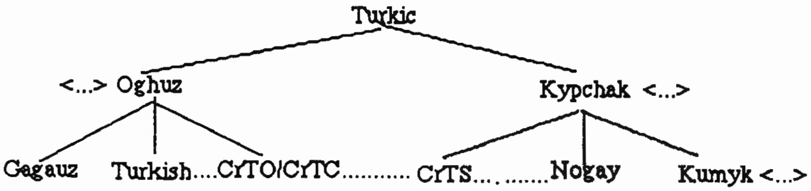


FIGURE 1.

While the major bulk of the entries in the word-lists represent Turkic words regardless of the groupings, some elements can be identified as typically Oghuz or Kypchak.

The words for 'dog' reflect a distinct Oghuz-Kypchak contrast:

Oghuz **köpek* 'dog' (CrTC *köpek,* CrTO *kȯpek′*, Kr *kopek)* — Kypchak **jit* (CrTS *it,* Nog *ijt).*

Oghuz also demonstrates some Persian loanwords, even in the basic word-list:

**tuqum* 'seed' < OPers **toxm* (OT *tuqum,* CrTC *tuqum,* CrTO *toqum)* — Kypchak **urluq* (CrTS *urlɨq,* Kr *urlux,* Nog *urlɨk);*

**qɨrmɨz* 'red' < OPers. **qeremez* (MPers. *qermez;* OT *kɨrmɨzɨ,* CrTO *qɨrmɨzɨ,* Kr *qɨrmɨzɨ)* — Kypchak **qɨzɨl* (CrTS *kɨzɨl,* CrTC *kɨzɨl,* Nog *kɨzɨl);*

**ateš* 'fire' < OPers **a:taš* (MPers *ateš;* OT *ateš,* CrTC, CrTO *ateš,* Kr *ateš)* — Kypchak **ot* (CrTS *ot,* Kar *ot,* Nog *ot).*

A typical Kypchak feature is the merging of the words for 'hand' and 'arm': Kypchak **qol* 'hand, arm' — Oghuz **el* 'hand' and **qol* 'arm'; cf. CrTS *qol* 'hand, arm,' CrTO and CrTC *qol* 'arm,' *el* 'hand.' Other Kypchak words:

**derek* 'tree' (CrTS *derek,* Kr *terek)* — Oghuz **aghač* (CrTC *aghač,* Kar *aghač,* OT *aghač);*

**jalda=* 'swim' < **jal* 'horse mane' [orig.: "to cross water holding on to the horse mane"][2] (CrT, Kr, Kar *jaldamaq)* — Oghuz *yüz=* (OT *yüzmek).*

In Kypchak, the prevalent word for 'bite' is the denominal verb **tišle=* (< **tiš* 'tooth'); while Oghuz also has the denominal verb (cf. OT *dišlemek),* a non-denominal form is more frequent (OT *isɨrmak* 'bite').

3. Grammatical elements. According to conventional wisdom, the lexicostatistical approach is not sufficient in itself to establish definite relationships between languages. Therefore, it is critical to check other elements of the language or to apply other analytical methods to the word-lists shown. Here the former option is taken; the result is a preliminary analysis of grammatical elements in CrT which bear resemblance to Oghuz or Kypchak. The present analysis will follow the description of Krymchak.

II. THE KRYMCHAK VARIANT: A GENERAL DESCRIPTION

1. Introduction. The Krymchaks of the Crimea, formerly a Turkic speaking people, represent a Jewish group primarily of Sephardic origin. The group used to practice Rabbinical Judaism, but the tradition is now entirely lost. In the relevant corpus of literature, the Krymchaks are either given no mention at all or are identified with the Karaites (Karaim) of the Crimea (for recent publications see Halen 1987:249). Some authors refer to the group very briefly, cf. the work by B.Pinkus (1988:6, 91, 266), who apparently uses secondary data.

By the 1979 estimate, Soviet Krymchaks number about 1,800 (Kupoveckij 1983:87). Major communities are found in Simferopol, Sukhumi and Novorossijsk. There are small Krymchak communities outside the USSR (see Pinkus 1984:87), however, they seem to have fallen under the influence of the neighboring population.

While earlier history of the Krymchaks is not quite clear (see Polinsky 1991), the group crystallized with the Sephardic influx to the Crimea, then an area of notable religious tolerance, and it expanded during the Crimean Khanate times (14th-16th centuries A.D.). The bulk of the Krymchak community were Spanish and

Portuguese Jews escaping religious persecution. The group also incorporated Italian Jews who arrived with the Genoese merchants[3]. After the Russian annexation of the Crimea in the late 1700s, the earlier elements merged with a certain number of European Jews from the north.

All the three elements mentioned are traced in Krymchak last names, cf. respectively Piastro, Portaqal/Portoqal, Manto, Xondo; Anželo; Ačkinazi, Varšavski. The general Sephardic element in the Krymchak community had, however, remained dominant. For more historical details see Filonenko 1972; Černin 1983 and Polinsky 1991. After the Russian Revolution and during the Civil War period, a number of the Krymchaks fled the Crimea, settling eventually in the Caucasus, Turkey, and Palestine. In the 1920s and early 1930s, Krymchak studies started in the USSR, and some publications appeared. Among these was a primer and two school-books edited by the notable Krymchak enlightener Isaac Kaja (Kaja 1928; 1930) and his colleagues (Čavtōbeli et al. 1933). These activities, however, were soon cut short by the Stalinist policies of Russification.

A great number of Krymchaks perished in the Holocaust, under the Nazi occupation of the Crimea, when they were identified with Ashkenazi Jews (Loewenthal 1951). Modern Krymchaks forcefully describe themselves as non-Jewish. As the result, their Krymchak identity is defined negatively: they focus on being different from (European) Jews, Karaites, Crimean Tatars and/or Russians.

Up to the 1930s, Krymchaks spoke a Turkic language[4] which has been classified either as a separate language (Shapshal 1916:IV,V; Filonenko 1972:9; Peysakh 1974; Moskowitch and Tukan 1986:118-121) or as a dialect of Crimean Tatar (Deinard 1878:115; Samoilovitch 1923:110; Kaja 1916; 1928; Černin 1983:98; 1988:167; Kozinskij and Polinskaja 1987; Polinskaja and Černin 1988:15). The language is currently almost extinct. It is remembered, though with a large degree of attrition, by speakers in their seventies[5].

2. Crimean Tatar and Krymchak. The same technique, as shown under I.2, was applied to classify Krymchak and Crimean Karaite, the language which has often been compared to it. Respective 100-word lists are shown in Appendix II. The results are presented in Table 2.

TABLE 2

Percentages of basic vocabulary: Crimean Tatar (Orta, Coastal, Steppe), Krymchak, Karaite

	Kar	CrTO	CrTC	CrTS
Kr	90%	89%	85%	83%
Kar		90%	86%	83%
CrTO			91%	82%
CrTC				77%

The data indicate that Krymchak and Crimean Karaite were closely related, probably as variants of one subdialect. Regular sound correspondences between the variants are shown in Table 3.

The Krymchak/Karaite subcluster is related to the Coastal/Orta Crimean Tatar cluster of which it is a subdialect[6]. From the viewpoint of linguistic taxonomy proper, the problem is thus solved[7]. However, there is another interesting problem here, namely that of the sociolinguistic status of Krymchak and Karaite. It is discussed below, in section IV.

TABLE 3.
Phonological correspondences between Krymchak and Crimean Karaite

Krymchak	Crimean Karaite	Example
1. word-final j	word-final z	*jaramaj* — *jaramaz* 'wicked,' but *qazmaq* 'dig'
2. intervocalic/postvocalic word-final:		
Ø	v	*suuq* — *suvuq* 'cold' *auz* — *avuz* 'mouth' *su* — *suv* 'water'
3. back rounded vowels	front rounded vowels	*butun* — *bütün* 'all' *olmek'* — *ölmek* 'die'

4. word-final *k:		
k'	k	*=mek'* — *=mek* (infinitival suffix)
5. word-final *q[8]		
x	q	*=max* — *=maq* (infinitival suffix) *qazax* — *qazaq* 'Russian' (> Russ. *kazak* 'cossack')
6. intervocalic gh	intervocalic v/w	*baghɨr* — *bavur* 'liver'
7. word-initial e/i before s/s:		*itmek'* — *ešitmek* 'hear'
Ø	e/i	

3. Linguistic description of Krymchak.

3.1. Phonology. The vocalic system of Krymchak is largely simplified as compared to Crimean Tatar. It consists of: *e a o u i*. As shown in Table 3, labialized vowels *ü* and *ö* are partially delabialized (respectively, to [u̇] and [ȯ]); they are positional variants of /u/ and /o/. The widespread Turkic opposition of /i/ and /ɨ/ is neutralized; the two vowels represent positional variants of one phoneme, /i/. The variant [i] occurs after the palatalized consonants and after palatal /č/, /š/ and /ǯ/ (cf. Table 4), e.g., *koriim* 'I see,' *kičkene* small,' *čiček'* 'flower,' *šitmek'* 'hear,' *ǯiğer* 'liver'[9]. Despite its palatal characteristics, /j/ only cooccurs with [ɨ], cf. *jɨl* 'year.'

One of the most striking features of Krymchak vocalism is the loss of vowel harmony. In fact, this is the continuation of the process started in Crimean Tatar proper (cf. Asanov, Garkavec and Useinov 1988:203). Here an attempt is made to relate this process to other changes in the phonological and accentuation system. In this respect, Krymchak proves to be particularly interesting, as it maximizes the tendencies already present in Crimean Tatar. It should, therefore, be borne in mind that the phenomena described here apply to Crimean Tatar, but to a lesser extent.

Besides the loss of vowel harmony, Krymchak treats the common Turkic rule of ultimate stress rather inconsistently. Cases are observed, as *´doğru* 'straight,' *tengrØ* < **tengrɨ* 'deity,' or *´vaxɨt* 'time' where the stress shifts away from the ultimate syllable.

Another feature of Krymchak is the development of a common Turkic tendency towards the reduction of unstressed vowels. Krymchak differs from other Turkic languages (which favor partial reduction: full vowel > schwa) in that it has a higher frequency of complete reduction, cf. the word for deity given above.

These seemingly unrelated phenomena can be explained in terms of a single vocalic hierarchy:

(3) a.. [i] > [e] > [a] > [o] > [u] > [ɨ]

which can be rewritten in a more traditional form as:

(3) b. front high V > front low V > back V[10].

The hierarchy predicts that the vowels to the left in (3a) will retain their properties (i.e., they will withstand reduction and harmonic assimilation). It also predicts that they will attract stress and control the vocalic properties of the environment (harmony) more efficiently than the vowels to the right in (3a), irrespective of their position in the word. Therefore, if all the vowels in a word are identical, there is no reason for a shift of stress from the final syllable, there is little reduction, and the harmony remains as a default rule, e.g., *pɨrnɨč* 'rice,' *tarağač/darağač* 'gallows.' Then if the word has different vowels, the stronger vowel is likely to attract stress and dissimilate from other vowels. Cf.:

(4) a. CrT bȯjun 'neck' — Kr ˈbojun 'neck' ([o] > [u])
b. CrT aša=j 'eats (3 sg.)' — Kr aši=j 'eats (3 sg.),'
aši=im 'eat (1 sg.)'/*aša=im ([i] > [a])
c. CrT tengri 'deity' — Kr tengr/*tngri 'deity'
d. CrT daˊvus/ˊdavus 'voice' — Kr ˊdavos 'sound'
e. OT faˊkir 'poor' — CrT 'fakir 'poor' — Kr f(ɨ)kaˊre
(> *fɨkare) 'poor'

It is important to note that (3a) represents a tendency, rather than a fixed rule; this tendency varies from speaker to speaker and is more prominent in modern Krymchak than in the traditional variant (for the latter see Text 1 below).

TABLE 4. KRYMCHAK CONSONANTS

	bilabial	dental	palatal	velar	uvular
occlusive	p	t	č	k, k'	q
	b, b'	d	ǯ	g	
fricative	f	s	š	x	
	v	z	j	ğ	
lateral	l, l′				
vibrant	r				
nasal	m	n		ŋ	

3.2. Morphology: Basic Data. The personal pronouns are shown in Table 5.

TABLE 5. KRYMCHAK PERSONAL PRONOUNS

	NUMBER	SG	PL.
PERSON	1	men	biz/bɨz
	2	sen	siz
	3	o(l)	o(lar)

The declension of 1. person sg. pronoun is as follows:

(5) Nom. men[11]
Gen. men=im
Dat. maŋa
Acc. men=i
Loc. men=de[12]

Demonstrative pronouns only distinguish between near and distant deixis: *bu* 'this' — *o* 'that.' The Crimean Tatar distinction between *su* 'this (speaker's deictic environment)' — *bu* 'this (hearer's deictic environment)' is lost.

The noun has the categories of number, case and possession. The possessive affixes are:

(6) 1sg =(i)m 1pl =miz
2sg =iŋ 2pl =ŋiz
3sg =(s)i 3pl =(s)i

The nominal declension is shown in (7) and (8).

(7) Declension: vocalic stem

(a) marked for possession *(ana=m* 'my mother'), (b) unmarked for possession *(Moskva* 'Moscow').

	(a)	(b)
Nom.	ana=m=Ø	Moskva=Ø
Gen.	ana=m=niŋ/ana=m=iŋ[13]	Moskva=ŋ
Dat.	ana=m=a/ana=m=ğa	Moskva=ğa
Acc.	ana=m=nɨ/ana=m=ɨ	Moskva=nɨ
Loc.	ana=m=da	Moskva=da
Abl.	ana=m=dan	Moskva=dan
Com./Instr	ana=m=nen	Moskva=nen

(8) Declension: consonantal stem

(a) marked for possession *(at=ɨm* 'my horse'), (b) unmarked for possession *(mɨšɨk* 'cat').

	(a)	(b)
Nom.	at=ɨm=Ø	mɨšɨk=Ø
Gen.	at=ɨm=ɨŋ	mɨšɨk=ɨŋ
Dat.	at=ɨm=ğa	mɨšɨk=qa
Acc.	at=ɨm=nɨ[14]	mɨšɨk=nɨ
Loc.	at=ɨm=da	mɨšɨk=ta
Abl.	at=ɨm=dan	mɨšɨk=tan
Com./Instr.	at=ɨm=nen	mɨšɨk=nen

The comparison between case forms in Crimean Tatar proper and Krymchak is given in Table 6.

The verb agrees with its subject in number and person. Agreement suffixes are presented in Table 7. The affixes vary according to the tense series, and the critical distinction is between the past and non-past tenses. Tense form markers are listed in Table 8.

TABLE 6. CASE ENDINGS: CRIMEAN TATAR AND KRYMCHAK

Case	CrT	Kr
Nominative	=Ø	=Ø
Genitive	=(i)ŋ/=niŋ[15]	=niŋ/=(i)ŋ
Dative	=ğa/=ge[16]	=ğa/=ge
	=qa/=ke[17]	=qa/=ke
	=(n)a/=(n)e[18]	=(n)a/=(n)e
Accusative	=ni	=(n)i
Locative	=da/=de	=da/=de
	=ta/=te	=ta/=te
	=nda/=nde	
Ablative	=dan/=den	=dan/=den
	=tan/=ten	=tan/=ten
	=nda/=nde	
Comitative	—	=nen

TABLE 7. AGREEMENT SUFFIXES IN CRIMEAN TATAR PROPER AND KRYMCHAK (PAST/NON-PAST)

Form	CrT	Kr
1 sg	=(i)m	=(i)m
2 sg	=ŋ/siŋ	=ŋ/=siŋ
3 sg	=Ø	=Ø
1 pl	=q/=miz	=x/=x:=miz[19]
2 pl	=ŋiz/=siŋiz	=siz/=siz:=(si)ŋiz[20]
3 pl	Ø=lar	Ø(=lar)

TABLE 8. TENSE MARKERS IN CRIMEAN TATAR PROPER AND KRYMCHAK (INDICATIVE)

Tense	CrTO/S	CrTS	Kr
Present Indefinite	=a=	=a=	=(a)i=
Non-Future	=r= + Aux	=maqta + Aux	—[21]

(Progressive/Aorist)			
Past Indefinite	=di/=ti	=di/=ti	=di/=ti
Perfect	—/=miš[22]	=ğan + Aux	=mišFuture
Definite	=aǯaq	=ǯaq	=aǯax

3.3. Syntax: Basic data. Like other Turkic languages, Krymchak is predominantly verb-final. As verb-finality is a particular case of the order in which the head follows the dependent element, Krymchak has modifier-modified (Adjective-Noun, Genitive-Nominal Possessor, Relative Clause-Head NP, Numeral-Noun) order, embedded clause preceding matrix, last name preceding first name. Krymchak has exclusively suffixes and postpositions.

However, deviations from the APV (Agent-Patient-Verb) order are quite frequent (cf. texts below), primarily towards AVP (SVO). The emergence of AVP as another important order may be explained by both internal factors related to language attrition, and the external influence of Russian which is predominantly verb-medial.

TABLE 9. BASIC NON-FINITE FORM MARKERS IN CRIMEAN TATAR PROPER AND KRYMCHAK

Non-finite form	CrTO/S	CrTS	Kr
Present Participle	=r	=a/=e[23]	=r=
(Subject Participle of simultaneity)	=araq	=araq	
Past Participle	=ip	=ip	(=ip)[24]
(Participle of priority)	=miš=	=ğan=	=miš=
Future Participle	=aǯaq	=aǯaq	=aǯaq
Specialized Object Participle	=DIğ=	---[25]	=DIğ=

The syntax of the complex sentence in Crimean Tatar is characterized by a heavy use of participial and gerundial constructions, typical of other Turkic languages. In the modern language, this syntax is gradually replaced by the conjunction of finite forms, both in Crimean Tatar proper and still more so in Krymchak. For the grammatical points made here, it is critical to look at the gerundial and participial affixes.

III. GRAMMATICAL DIFFERENCES BETWEEN THE OGHUZ AND KYPCHAK GROUPS AS REFLECTED IN CRIMEAN TATAR AND KRYMCHAK

In the nominal morphology, the combination of Oghuz and Kypchak features is reflected in the heterogeneity of the nominal paradigm (cf. (7), (8), and Table 6). The consonant-initial Accusative and Dative affixes are typically Kypchak (cf. Acc. *=i/ü* and Dat. *=a/=e* in OT); the Ablative forms are more Oghuz-like. The Genitive forms split according to the distinction between CrTO and CrTC on the one hand, and CrTS on the other. The former use the Oghuz type Genitive without a "buffer" *n,* the latter has the Kypchak "buffered" form *=niŋ*. The same Kypchak form is used in Krymchak. The Instrumental form is of the Kypchak origin, though a number of modern Kypchak languages do not have it, e.g., Kazan Tatar.

The verbal paradigm seems clearer with regard to variant distribution. The Future tense is marked by the typically Oghuz form in *=VǯVq* (Definite Future/Necessitive). The participial nature of the latter form is still retained, cf. Table 9 and examples like (9) in Text 1 below:

(9) doğr=aǯax qisajax=lar
'pregnant women (lit.: who are due)'

The Perfect form in *=miš* is exclusively Oghuz, and is observed only in the Orta and Coastal variant. In addition to this form, these two variants of Crimean Tatar and Krymchak have the Object Participle in *=DIğ=* which is also exclusively Oghuz. On the other hand, the language has the participle of priority (Past Participle) in *=ip* which is typically Kypchak (cf. Džanmavov 1967:33-67). The Steppe variant has the Kypchak participle in *=ğan=* which is absent from Coastal and Orta Crimean Tatar.

The 1 person plural agreement marker *=q/=k* is characteristic of the Oghuz languages, cf. Ottoman Turkish in (10). Parallel to it, the common Turkic marker *=miz* is used in all the variants.

(10) bak=tɨ=k
look=PAST=1PL
'we looked'

An interesting case is the Krymchak Present Tense suffix *=ai=/=i=* which is definitely non-Oghuz and does not have straight formal parallels in Kypchak languages, cf.

(11) kel=i=Ø=ler
come=PRES=3SG=PL
'they come, are coming'

The origin of this marker can be tentatively explained by the reanalysis of a non-finite form. Indeed, Kypchak languages have the Simultaneity (Present) Participle in _*=i/=(*i)V_[26], cf. in Karachay (12a) and Crimean Tatar Steppe (12b):

(12) a. tübe=j
meet=PART
'meeting, when meeting'
b. kel′=e
come=PART
'arriving'

The reanalysis of this participle into a finite form (Present Tense form) is attested in Kazan Tatar, cf. the Present in *=j* (*=j* affixes to the stems ending in vowel, while *=a/=i̱* are the harmonic variants attached by consonantal stems):

(13) Bolɨt=lar kük=ne kaplɨ=j=lar
cloud=PL sky=ACC cover=PRES=PL

Thus, the structural formation is consistent with the data from Kypchak languages. The distinct form of the suffix can be explained by the fact that, after the participle-to-finite verb reanalysis, the paradigm was patterned after the verbs ending in a vowel. The verbs with the back vowel(s) in the stem started to attach the =ai= suffix; the verbs with the front vowel(s) in the stem attached the *=ei= suffix. Since Krymchak [i] and [e] are quite close, the reanalysis **=ei=* > **=ii=* > *=i=* can be posited. The intermediate form is still attested in Krymchak schoolbooks (Kaja 1928; 1930), cf.

(14) kel=ij=Ø
come=PRES=3SG
'he comes'

To summarize, the grammatical structure of Crimean Tatar reveals a mixture of the two genetic groups but is consistent with the classification based on vocabulary items. The verbal morphology, in particular, confirms the lexicostatistic classification and indicates the heterogeneity of what is conventionally considered the Crimean Tatar language[27]. The Coastal and Orta variants prove to be closer to the Oghuz group, while Steppe Crimean Tatar has more Kypchak features. As the grammatical system is less open to contact influence, it seems plausible to assume the classification proposed above.

The status of Krymchak poses other questions: first, Krymchak demonstrates a greater mixture of the Oghuz and Kypchak features; second, there is more than average variation across idiolects and even within idiolects (cf. Texts 2 and 3 below). This problem is addressed in the next section.

IV. THE NOTION OF ETHNOLECT

The genetic classification notwithstanding, the problem remains of fitting Krymchak into a sociolinguistic taxonomy. It becomes still more intriguing in view of the fact that Krymchak, which has always been spoken by a small language community[28], demonstrates exceptional variation between local versions (e.g., that of Simferopol, Feodossia, Qarasubazar) and even between individual speakers. Speakers are well aware of it and insist on local differences.

Some examples of the variation in Krymchak were demonstrated above, in the discussion of morphology; cf. also *učine* and *ičine=ge* both meaning ‘inside’ in Text 2 and 3, respectively. One could expect such a variation in a major language spoken in large territories and by large communities, not in a small subdialect. This variation could be tentatively explained by the current language attrition, but it is also found in earlier texts (Polinsky 1991; Shapshal 1916:IV; Deinard 1878:115-116).

In addition to this variation, Krymchak demonstrates a greater mixture of the Oghuz and Kypchak features discussed above, than any of the three tentative Crimean Tatar versions.

This diversity seems to be motivated by sociolinguistic, rather than structural factors.

Though the early history of the Krymchaks is not quite clear, it is apparent that they were located in urban communities and, by the 1500s[29], shared the Turkic language of the neighboring population. The history of the attested Krymchak centers shows that they gradually moved from city to city, in particular inland from the coastal areas. Thus, in the 1700s, they were primarily settled in Esqɨ Qɨrɨm; by the early 1800s, the center moved to Qarasubazar, and by the 1910s, to Simferopol (Černin 1983:96; Kupoveckij 1983:86-88). Krymchak, therefore, developed in the intelligible Crimean Tatar environment; as the environment changed slightly, these changes led to a greater heterogeneity of Krymchak. The diversity increased in the urban setting which was favorable for the union of different vernaculars. Thus, Z. Deinard (1878:116) identifies Krymchak with an urban jargon. The distinct nature of Krymchak as an idiom was reinforced by the religious difference between the Jewish speakers and their Moslem environment.

While it seems feasible to reserve the notion of jargon for the case where there are at least two genetically unrelated contact languages, the Krymchak situation can be adequately described as that of ethnolect[30].

Ethnolect is defined as the variant of the single or predominant language as it is spoken by a distinct and non-assimilated minor group within the dominant population. For this notion it irrelevant whether the minor group has or had another language. To put it informally, the essence of ethnolect is a compromise between linguistic assimilation and cultural, ethnic and/or religious dissimilation.

The relationships between the standard predominant language and the ethnolect can be characterized according to the degree of structural differentiation. The ethnolect may differ just slightly from the standard, or it may be a subdialect or a dialect of it. Thus, the notion is categorically different from that of the (sub)dialect.

The use of this notion seems quite promising in the hierarchical multilingual society of the Soviet type, where much language variation can be found. In particular, this notion can be applied to the variants of Russian spoken in ethnic republics or areas of the Soviet Union.

CONCLUSIONS

The conventional classification of Turkic languages of the Crimea has been analyzed and questioned. An alternative classification was suggested which distinguishes between the two variants of Crimean Tatar (Orta and Coastal Crimean Tatar) on the one hand, and Steppe Crimean Tatar on the other. While the former variants seem related to the Oghuz grouping of Turkic, the latter belongs to the Kypchak grouping. Thus, the homogeneity of the accepted Crimean Tatar language proves problematic.

In the case of closely related languages, the classification should rely on both lexical correspondences and correspondences between relational morphemes. All the reservations against morphological innovations notwithstanding, the latter criterion is credible for closely related languages, as the time of the possible change is quite short. In the Crimean Tatar case, the lexicostatistic and morphological data prove consistent.

Krymchak, a subdialectal variant of Orta/Coastal Crimean Tatar, formerly spoken by the Sephardic Jewish group of the Crimea, was briefly described. Krymchak differs from Crimean Tatar proper in that it has more of a mixture of the Oghuz and Kypchak elements. It also shows greater variation across local variants and/or idiosyncrasies of individual speakers. This can be due to sociolinguistic factors. To describe the variation between Crimean Tatar and Krymchak, the notion of ethnolect was proposed.

APPENDIX I. BASIC WORD LISTS FOR THE THREE TENTATIVE DIALECTS OF CRIMEAN TATAR.

WORD	CrTO	CrTC	CrTS
all	butu̇n	bütün	butun
ashes	ku̇l′	kül	kül
bark	qabuq	qabuq	qabɨq
belly	qursaq	qarɨn	xursax
bird	quš	quš	quš
big	bujuk	büjük	bijk
bite	tišle=	čšle=	tišle=
black	qara	qara	qara
blood	qan	qan	qan
bone	süjek	süjek	kemik
breast	kokus	göküs	koks
burn	jan=	ǯan=	ǯan=
claw	tɨrnaq	tɨrnaq	tɨrnax
cloud	bulut	bulut	bulut
cold	suvuq	suuq	suux
come	kel′=	gel′=	kel′=
die	ȯl′=	öl=	ol′=
dog	kopek′	köpek	it
drink	ič=	ɨš	ič=
dry	quru	quru	xatɨ
ear	qulaq	qulaq	qulax
earth	jer	topraq	jer
eat	aša=	aša=	ǯe=
egg	jɨmɨrta	jɨmɨrta	jumurta
eye	koz	göz	köz
fat	jağ	jav	maj
feather	quš=qanat	lelek	kušɨn
fire	ateš	ateš	ot
fish	balɨq	balɨq	balɨx
fly	uč=	üč=	ič=
foot	ajaq	aǯaq	taban
full	tolu	dolu	tolɨ
give	ber=	ver=	ber=
good	jaxšɨ	jaxšɨ	egije

green	ješli	ješli	ǯašlɨ
hair	sač	sač	sač
hand	el	el	qol
head	baš	baš	baš
hear	ešit=	išit=	išit=
heart	jürek	jürek	jürek′
horn	mojuz	muguz	mujuz
I	men	men	ben
kill	ȯldur=	öldür=	ȯldur=
knee	tiz	diz	tiz
know	bil′=	bil′=	bil′=
leaf	japraq	ǯapraq	ǯaprax
lie	jat=	ǯat=	ǯat=
liver	qara=ǯiger	ǯiğer	bavɨr
long	uzun	uzun	uzun
louse	bit	bit	bit
man	er	marda	erkek
many	čok	čok	köp
meat	et	et	et
moon	aj	aj	aj
mountain	dağ	tau	bair
mouth	ağɨz	ağɨz	ağɨz
name	ad	ad	ad
neck	bojun	moin	boin
new	eŋɨ	jenɨ	ǯana
night	geǯe	geǯe	keǯe
nose	burun	burun	burun
not	=mV=	=mV=	=mV=
one	bir	bɨr	bɨr
person	adam	adam	kiši
rain	jağmur	ǯağmur	ǯağmur
red	qɨrmɨzɨ	qɨzɨl	kɨzɨl
road	jol	jol	ǯol
root	tamɨr	tamur	tamɨr
round	togerek	tögerek	tȯgerek
sand	qum	qum	xum
say	sȯjle=	söjle=	ajt=
see	kor=	gör=	kur=
seed	urluq	tuqum	urlɨk
sit	otur=	otur=	otur=
skin	tere	dere	teri
sleep	juxla=	juxla=	ǯat=

small	jufaq	jufaq	kišik
smoke	tutun	duman	duman
stand	tur=	tur=	tur=
star	jɨldɨz	jɨldɨz	jɨldɨz
stone	taš	daš	taš
sun	ku̇neš	güneš	kuneš
swim	jalda=	yüz=	ǯalda=
tail	qujruq	qurjuq	qurjux
that	o	o	o/a
this	bu	bu	bu
thou	sen	sen	sen
tongue	til′	dil′	til′
tooth	tɨš	čiš	tɨš
tree	terek	ağač	derek
two	eki	iki	iki
walk	jur=	bar=	ǯur=
warm	čɨllɨ	jilli	sɨǯax
water	suv	su	su
we	biz	biz	bɨz
what	ne	ne	ne
white	bijaz	bijaz	ax
who	kim	kim	kim
woman	qadɨn	qadɨn	xɨsajaxlɨ
yellow	sarɨ	sarɨ	sarɨ

APPENDIX II. BASIC WORD LISTS FOR KRYMACHAK AND CRIMEAN KARAITE

WORD	Kr	Kar
all	epsi	bütun
ashes	kül′	kül′
bark	qabuq	qabuq
belly	qursaq	qursaq
bird	quš	quš
big	balaban	balaban
bite	tɨšle=	čɨšle=
black	qara	qara
blood	qan	qan
bone	kemik’	kemik
breast	kokus	emček

burn	jan=	jan=
claw	tɨrnaq	tɨrnaq
cloud	bulut	bulut
cold	suuq	suvuq
come	kel′=	kel′=
die	ȯl′=	öl=
dog	kopek′	köpek
drink	ič=	ɨč
dry	quru	quru
ear	qulaq	qulaq
earth	jer	jer
eat	aša=	aša=
egg	jɨmɨrta	jɨmɨrta
eye	koz	köz
fat	jağ	jav
feather	pux	qanat
fire	ateš	ateš
fish	balɨq	balɨq
fly	uč=	üč=
foot	ajax	taban
full	tolu	tolu
give	ver=/ber=	ber=
good	jaxšɨ	jij
green	ješli	ješil′
hair	sač	seč
hand	qol	qol
head	baš	baš
hear	ešit=	ešit=
heart	jürek	jürek′
horn	mojuz	muguz
I	men	men
kill	ȯldur=	öldir=
knee	tiz	tɨz
know	bil′=	bil′=
leaf	japraq	japraq
lie	jat=	jat=
liver	bağɨr	bavur
long	uzun	uzun
louse	bit	bit
man	erkek	erkek
many	čok	čok
meat	et	et

moon	aj	aj
mountain	dağ	qɨr
mouth	ağɨz	avuz
name	ad	ad
neck	boin	bojun
new	jeŋɨ	jenɨ
night	keče/geǯe	geǯe
nose	burun	burun
not	=mV=	=mV=
one	bir	bɨr
person	adam	adam
rain	jağmɨr	jağmur
red	qɨrmɨzɨ	qɨrmɨzɨ
road	jol	jöl
root	tamɨr	tamɨr
round	togerek	tögerek
sand	qum	qum
say	sȯjle=	söjle=
see	kor=	kör=
seed	tuqum	urluq
sit	otur=	otur=
skin	terɨ	tere
sleep	juxla=	juxla=
small	kičkene	jufaq
smoke	tutun	tutun
stand	tur=	tur=
star	jɨldɨz	jɨldɨz
stone	taš	taš
sun	ku̇neš	küneš
swim	jalda=	jalda=
tail	qurjux	qujruq
that	o	o
this	bu	bu
thou	sen	sen
tongue	tɨl′	til′
tooth	tɨš	čiš
tree	terek	ağač
two	eki	iki
walk	var=/bar=	bar=
warm	sɨǯax	jillɨ
water	su	suv
we	biz	bɨz

what	ne	ne
white	bijaz	aq
who	kim	kim
woman	qadɨn	qadɨn
yellow	sarɨ	sarɨ

APPENDIX III. KRYMCHAK TEXT SAMPLES.

A large corpus of Krymchak texts is given in Polinsky 1991. Here, a few short texts and an excerpt are presented which support the linguistic facts discussed above. All the texts are given in phonological transcription. Reduced sounds are put in parentheses.

Text 1. This text is an excerpt from the Passover Haggadah (Kr *ağada)*. The text was transcribed from a recording by Yevsey Peysakh (ca. 1972)[31]. It is unclear what written text Peysakh was using. A published Haggadah *(Hağada šel petax* [sic!]) is a likely option; however, its text differs slightly from Peysakh's presentation and has more Crimean Tatar idiosyncratic features. Another option could be a hand written home journal *(ǯonka);* such journals were used by the Krymchaks until the early 1930s, and Peysakh, who was a scrupulous collector, had several *ǯonkas*[32].

B′il′=e=sin oğlu=m ke o vax(i)t=ta Nimrud
know=OPT=2SG.OPT son=my PRT that time=LOC Nimrod

padiša=niŋ, kunu=nde ki o bujuk ve küčlü padiša=Ø
king=GEN day=LOC PRT that great and mighty king=NOM

e=di=Ø, put-vurgun e=di=Ø=ler o vax(i)t=ni
be=PAST=3SG idol-worshipper be=PAST=3SG=PL that time=ACC

xalx=Ø. Ve padiša=si bujuk=le=n=ir e=di=Ø.
people=NOM and king=3SG.POSS great=VBLZ=n=PART be=PAST=3SG

De=r e=di=Ø ke: men jarat=miš=im dunja=ni kök=nü
say=PART be=PAST=3SG PRT I create=PERF=1SG world=ACC sky=ACC

ve jer=ni ve ǯumle max(u)qit=ler=ni men jarat=miš=im.
and earth=ACC and all surrounding=PL=ACC I create=PERF=1SG

Xalx=ta, oŋa inan=ip, oŋi tengri beli=ip, baš
people=and he(DAT) believe=PART he(ACC) deity know=PART head

ur=ar e=di=Ø=ler oŋa <...> Nimrud padiša=ğa bend=PART be=PAST=3SG=PL he(DAT). N king=DAT

ağa=lar=i de=di=Ø=ler ki bir erkek bala
adviser=PL=3SG.POSS say=PAST=3SG=PL PRT one man child

doğ=aǯax=tir dunja=da. O bala seniŋ
be born=FUT=PRED world=LOC that child thy

padiša=lix=in=ni bujuk=lix=in=ni iŋka=r
king=NMLZ=POSS.2SG=ACC great=NMLZ=POSS.2SG=ACC replace=PART

et=ter=eǯek=dir dunja=dan. Qačan ki sȯz=u̇=ni
do=CAUS=FUT=PRED world=ABL when PRT word=POSS.3SG=ACC

sit=ti=Ø, padiša=Ø šaat=i=nda eniṛ
hear=PAST=3SG king=NOM hour=POSS.3SG=LOC decree

čix=ir=di=Ø ki doğr=aǯax qisajax=Ø=lar
go out=CAUS=PAST=3SG PRT bear=FUT woman=NOM=PL

var ol=maja ki gendi=ler=na ev=ler=i=nde
be(IRREAL) be=VN PRT self=PL=DAT house=PL=POSS.3SG=LOC

doğur=ma=sin=Ø=lar.
bear=NEG=OPT=3SG=PL

'You should know, my son, that at the time of king Nimrod, who was a great and mighty ruler, at that time people worshipped idols. And this king waxed proud. He used to say: I created the world, and the heavens, and the earth, and I created all our surroundings. And the people believed in him and took him for their god, and worshipped him.<...> Nimrod's courtiers told him that a child would be born into the world who will make (Nimrod's) kingship and greatness denied in this world. When Nimrod heard these words, at the same moment he decreed that any woman who was due at that time could not bear the child at her home.'

Text 2. Unlike Text 1, this text uses everyday vocabulary and structures. This is a recipe, which was recorded by the author in Simferopol, 1987, from the 77 year old David Gabaj who at that time was one of the most fluent Krymchak speakers. The recipe describes small mutton dumplings, similar to tortellini — *suzme* (< Common Turkic **süz* = 'strain'—when ready, the dumplings are set on a colander).

xamur=Ø basla=ǯa=sin, suv=nen, tuz=nen.

dough=ACC.INDEF fold=FUT=2SG water=INSTR salt=INSTR

xamur=ni jaj=aǯa=sin, fa(x)=čix-fa(x)=čix, togere(k)=čix.
dough=ACC roll=FUT=2SG small=DIM round=DIM

učine et=Ø qoj=aǯa=sin, soŋ qapa=t=aǯa=siz
inside meat=ACC.INDEF put=FUT=2SG after cover=CAUS=FUT=2PL

baš=čix=iŋ. ol=ai=Ø suzme. učne jağ
head=DIM=POSS.2SG be=PRES=3SG suzme for fat

qoj=aǯa=sin, toq=ai=sin, jağ=la=Ø=sin, siǯax
put=FUT=2SG pour=PRES=2SG fat=VBLZ=OPT=OPT.3SG hot

ol=Ø=sin.
be=OPT=OPT.3SG

'Prepare dough with water and salt. Roll out the dough, cut out small round puffs. Put ground meat inside each puff, and close the dumpling's head. This is suzme. To put fat on it, pour it, so that it gets covered with fat. The fat should be hot.'

Text 3. This is another recipe, told to the author in Simferopol, 1987, by two women in their early seventies (Sara Gabaj and Simxa Čolaq). The text illustrates the dramatic degree of language attrition and actually sounds like a text in a pidgin of a kind. The interfering language is Russian. Russian words are underlined. The recipe describes one of the most important items in the Crimean Tatar and Krymchak cuisine, *kubete*. It is a large round meat pie; the pie is closed, and the cover used to be decorated by a strikingly intricate ornament.

kubete nazyvajetsja. unne=ni ele=me=ge kerek.
kubete is called flour=ACC sift=VN=DAT necessary

potom vzjat' bir mišlepe su tuz nemnožečko, peremešat',
then to take one cup water salt a little to mix

tuz=li suv=čix sdelat. ičine=ge qoj=ma=ga. potom značit
salt=ADJ water=DIM to make inside=DAT put=VN=DAT then so

suv=nen xamur basla=ma=ğa. tak vot xamur=Ø basla=i=Ø,
water=INSTR dough knead=VN=DAT so dough=ACC knead=PRES=3SG

a potom èto xamur=Ø ostavljajete, tašla=i=siz, čtob

and then	this	dough=ACC	leave(2PL,PRES)	leave=PRES=2PL	so that	

qapi=la=ǯa	tur=ai=Ø	bir=bača.	a	potom
cover=VBLZ=FUT.PART	stand=PRES=3SG	some time	and then	

al=ai=siz,	kačalka=nen	katajete.
take=PRES=2PL	rolling pin=INSTR	roll out(2 PL,PRES)

'The dish is called *kubete*. Sift the flour, take one cup water and some salt. Mix the water and salt together, pour the salty water into the flour. Knead thoroughly and fold. Leave the dough covered. Roll out with a rolling pin.'

REFERENCES

Asanov Š. A and A.N.Garkavec, S.M.Useinov. 1988. Krymskotatarsko-russkij slovar'. Kiev: Radjans'ka škola.

Čavtōbeli Z.; R.Mullina; and S.Abdurahimof. 1933. Qirimčaq ibtidaji mekteblerine maxsus elifbeden soŋ oxuv kitabɨ. Simferopol: Krymgosizdat.

Černin V.Ju. 1983. O pojavlenii ètnonima "krymčak" i ponjatija "krymčakskij jazyk." In: Geografija i kul'tura ètnografičeskix grupp tatar v SSSR, ed. I.I.Krupnik. Pp.93-104. Moscow: Geografičeskoe obščestvo.

___. 1988. Krymčakskoe proiznošenie ivrita. In Sinxronija i diaxronija v lingvističeskix issledovanijax. Pt.2, ed. V.I.Podlesskaja. Pp.166-174. Moscow: Nauka.

Deinard Z. 1878. Massa krim. [Warsaw]: n.p.

Deny J., ed. 1959. Philologiae Turcica Fundamenta. I. Wiesbaden: Franz Steiner.

Doerfer G. 1959. Das Krimtatarische. In Deny 1959. Pp.369-390.

Džanmavov Ju.D. 1967. Deepričastija v kumykskom literaturnom jazyke. Moscow: Nauka.

Filonenko V.I. 1972. Krymčakskie ètjudy. Rocznik orientalistyczny 35:5-35.

___ and I.S.Kaja. 1955. Krymčaki i ix narodnoe tvorčestvo. Odessa. MS.

Hağada šel petax [sic!]. 1904. Petrokov: Zederbaum.

Halen H. 1987. Handbook of oriental collections in Finland. London: Curzon.

Hesche W. and H.Scheinhardt. 1974. Eine krimtatarische Wörterliste. Central Asiatic Journal 18:227-250.

Kaja I.S. 1916. Krymčaki. Evrejskaja starina 9:15-26.

___. 1928. Qrɨmčaq balalarɨ ičyn ana tilinde alefbet ve oquv kitabɨ. [Simferopol]: Qrɨm devlet nešriatɨ.

___. 1930. Oquv kitabɨ. Qrɨmčaq mekteblerinin ekinči sɨnɨfna maxsus. [Simferopol]: Qrɨm devlet nešriatɨ.

Kozinskij, I.S. and M.S.Polinskaja. 1987. Ob odnom slučae kontaktnogo vzaimovlijanija rodstvennyx jazykov. In Vozniknovenie i funkcionirovanie kontaktnyx jazykov, ed. I.F.Vardul and V.Belikov. Pp.56-59. Moscow: Nauka.

Kozinsky I.S. and M.S.Polinsky (to appear). Krymchak phonology.

Kupoveckij M.S. 1983. Dinamika čislennosti i rasslenie karaimov i krymčakov za poslednije 200 let. In Geografija i kul'tura ètnografičeskix grupp tatar v SSSR, ed. I.I.Krupnik. Pp.75-93. Moscow: Geografičeskoe obščestvo.

Loewenthal R. 1951. The extinction of Krimchaks in World War II. The American Slavic and East European Review 10:130-136.

Malowist M. 1947. Kaffa — kolonia genueńska na Krymie i problem wschodni w latach 1453-1475. Warsaw: Ossolyneum.

Moskowitch W. and B.Tukan. 1986. Quelques traits caractéristiques du parler des juifs krimtchaks de Crimée. Massorot 2:113-122.

Peysakh Y. I. 1974. Krimtshakhes. Sovetish heymland 9:138-144.

Pinkus B. 1984. The Soviet government and the Jews, 1948-1967: a documented study. Cambridge: Cambridge University Press.

__.1988. The Jews of the Soviet Union: the history of a national minority. Cambridge: Cambridge University Press.

Polinskaja M.S. 1988. Krymčaki: k istorii i tolkovaniju ètnonima. In Problemy jazykovogo var'irovanija i normirovanija, ed. N.I.Filičeva and T.Mixailova. Pp.45-48. Moscow: Moscow University Press.

__ and V.Ju.Černin. 1988. Terminologija rodstva u krymčakov. Sovetskaja tjurkologija 3:15-23.

Polinsky, M. 1991. The Krimchaks: history and texts. Ural-Altaische Jahrbücher 63:137-172.

Samoilovič, A.N. 1923. Nazvanije dnej nedeli u tureckix narodov. Jafetičeskij sbornik 2:98-118.

Sevortjan È.V. 1966. Krymsko-tatarskij jazyk. In Jazyki narodov SSSR, 1. Tjurkskije jazyki, ed. N.A.Baskakov. Pp.234-259. Moscow: Nauka.

Šapšal S.M. 1916. Skazka na jazyke krymskix evreev (krymčakov), zapisannaja v Karasubazare letom 1911 g. Zapiski vostočnogo otdelenija Imperatorskogo russkogo arxeologičeskogo obščestva. [Protokoly] 23 (Fasc.III-IV):IV-V.

Voegelin C. and F.Voegelin. 1977. Classification and index of world's languages. Amsterdam: Elsevier.

NOTES

[1] The data on Krymchak represent a fragment of a larger fieldwork project which was started by the author in collaboration with Isaac Kozinsky and Velvl Chernin, whose cooperation is gratefully acknowledged here. The data were obtained by the author during fieldwork in Simferopol (1987 and 1988), Moscow (1986-1988) and Leningrad (1987), and by Isaac Kozinsky in Sukhumi in 1987 and in Moscow 1987-1988. (For a detailed description of the speakers see Polinskaja and Černin 1988:16 and Polinsky 1991). For this study, 32 Krymchak speakers were consulted. Crimean Tatar speakers were studied by the author in Simferopol (1987),

Moscow (1987-1988) and Leningrad (1988). The sample included speakers of all the three dialects. If not specified otherwise, the data from the Orta dialect, which is the basis of the standard language, are taken as the point of departure.

The following abbreviations are used: CrT — Crimean Tatar; CrtC — Crimean Tatar, Coastal dialect; CrTO — Crimean Tatar, Orta dialect; CrTS — Crimean Tatar, Steppe dialect; Kar — (Crimean) Karaite; Kr — Krymchak; Nog — Nogay; OT — Ottoman Turkish. Abbreviations used in the interlinear glosses: ACC — Accusative; ADJ — Adjective, CAUS — Causative, DAT — Dative, DIM — Diminutive, FUT — Future, GEN — Genitive, INDEF — Indefinite, INSTR — Instrumental, LOC — Locative, NMLZ — Nominalizer, NOM — Nominative, OPT — Optative, PART — Participle, PERF — Perfect, POSS — Possessive, PRED — Predicative, PRT — Particle, VBLZ — Verbalizer, VN — Verbal Noun.

[2] This etymology is quite transparent in view of the nomadic heritage of the Kypchaks (L.S.Levitskaya, personal communication).

[3] Genoese merchants had strong colonies in the Crimea until the late 1400s, cf. Malowist 1947.

[4] Hebrew, which very few speakers read or understood, was retained for religious services. On the Krymchak version of Hebrew, which is close to the Sephardic variant, see Černin 1988.

[5] While the extinction of Krymchak can be motivated, among other reasons, by the small number of speakers, similar processes of attrition, though less pronounced, are observed for Crimean Tatar proper (meanwhile, the Crimean Tatar group is estimated at about one million — V. Shemizadeh, personal communication).

[6] An interesting problem would be comparing the Karaite dialects of Lithuania and the Ukraine/Poland to the Crimean variant. This, however, is beyond the points made in this paper.

[7] There seems little reason in V. I. Filonenko's (1972: 7) classification of Krymchak as Chagatay. First, the Krymchaks themselves have never been aware of the term and have never applied it to their language. Chagatay is the denomination of the late medieval *literary* language which was used in Samarkand and Andizhan; the classical period of Chagatay is associated with Alisher Navoi's writings.

[8] In this feature, Krymchak differs most conspicuously from Crimean Tatar; Karaite retains the Crimean Tatar consonant.

[9] An exception to the [ǯ + i] rule is the word *Axmeǯit* 'Simferopol' which retains the Crimean Tatar pronunciation.

[10] The hierarchy was first suggested for the description of Krymchak stress by I. S. Kozinsky (personal communication) and is analyzed in greater detail in a joint paper (Kozinsky and Polinsky, to appear).

[11] The variant *ben* is also attested, though less frequently. Most speakers allow for only one variant, however, some use both variants without any consistent distribution rule.

[12] Crimean Tatar also has the Ablative form *(menden)* which is not attested in Krymchak.

[13] Here and elsewhere the forms divided by a slash represent the variants obtained from different speakers. The discussion of such variation is given in section IV.

[14] One speaker out of 32 studied also suggested *at=ɨm=ɨ*..

[15] The form *=iŋ* is attested in the Coastal and Orta variants, the form *=niŋ* is Steppe Crimean Tatar.

[16] With consonantal stems ending in a voiced consonant other than nasal.

[17] With consonantal stems ending in a voiceless consonant.

[18] Selected by several stem types: (a) consonantal stems ending in a nasal, which include stems marked for possession; (b) vocalic stems, which include stems marked for possession by a suffix ending in a vowel. Stems (a) select the *=a/=e* variant.

[19] There seems to be variation of the two affixes across idiolects and even within one idiolect.

[20] See note 18.

[21] While absent from modern Krymchak, this form was apparently in use earlier, cf. *de=r edi* 'they said' in Text 1. It is, however, unclear whether it was a genuine Krymchak form or a conscious use of the CrT tense.

[22] Only in the Coastal dialect, where both the participial and the finite forms in *=miš* are observed.

[23] See (12b) below.

[24] Cf. note 20.

[25] No specialized form is attested; the function is performed by the participle of priority in *=ğan*, cf. *ǯazil=ğan mektüp* 'the letter written' (Object Participle) — *keč=ken afta* 'last week (the week which is gone).'

[26] The ending =j attaches to the stems ending in a vowel; the stems ending in a consonant attach =a/=e, according to the vowel harmony rule.

[27] Incidentally, this heterogeneity is indirectly reflected by speakers' intuitions. The speakers of Coastal and Orta Crimean Tatar consider their idioms to be closely related, while the Steppe variant is estimated as a totally different language (and vice versa). On the other hand, the subtlety of the distinctions is confirmed by the fact that younger speakers in the Soviet Union (those who grew up away from the dialectal division and were heavily exposed to Russian and/or Uzbek) and the majority of speakers outside the Soviet Union (e.g., in Romania, USA) do not quite differentiate between the variants.

[28] The maximum estimate, before World War II, was about six thousand speakers (Filonenko and Kaja 1955).

[29] This was the time when they were first mentioned in the decrees of Crimean Khans, cf. Černin 1983.

[30] The term was suggested by V.Černin (1983: 97-98; 1988: 167).

[31] Yevsey Peysakh was an amateur collector of Krymchak folklore, language materials, and items of material culture. In 1974 he published a paper on the Krymchaks in the Soviet Yiddish magazine *Sovetish heymland* (Peysakh 1974). He was also author of the article about the group in the *Sovetskaja ènciklopedija*, 3rd Edition. The tape recording is courtesy of Y.Peysakh's son, B.E.Peysakh.

[32] The tape was checked with Vildan Shemızadeh, a fluent speaker of the CrT Orta, Dialect who was only unable to identify Hebrew borrowings and some Arabic words no longer used in Crimean Tatar.

HOW CAN CLASS MARKERS PETRIFY? TOWARDS A FUNCTIONAL DIACHRONY OF MORPHOLOGICAL SUBSYSTEMS IN THE EAST CAUCASIAN LANGUAGES

Wolfgang Schulze-Fürhoff
(Paris/Bonn)

SUMMARY

Most of the East Caucasian languages (ECL) are characterized by a more or less elaborated system of noun classification, which is covertly encoded by means of so-called class markers (CM). Though the (internal and external) genetic background of ECL is far from being totally elucidated, many investigators assume a parallel technique at least for the hypothetical East Caucasian proto-language and try to derive as many inflectional morphemes as possible of the individual languages from this paradigm. The method is especially productive in the four ECL that do not exhibit a class marking strategy, proposing a process of petrification, i.e., lexicalization of once productive CM in these languages. In this paper I want to ask which functional conditions must be met so that a (reconstructed) class marker could have petrified. The study can be regarded as an attempt to help to fill the "diachronic gap" common to nearly all ECL by means of what is here called a "Functional Diachrony".

1. Introduction: When reviewing works on the reconstruction of Proto-East Caucasian (PEC), we observe that the assumption of so-called "petrified class markers" (PCM) plays an important role in examining verbal, pronominal, and nominal roots in order to establish word correspondences between the at least twenty nine autochthonous East Caucasian languages (ECL).

The basic problem when dealing with these languages diachronically is the fact that, with the exception of some sporadic findings, they do not have any relevant documented history. Most of the ECL are first documented by European researchers beginning in the 19th century; only for Avar, Dargwa, and Lak are some earlier sources (defectively written in the Arabic alphabet[1]) attested due to their semi-official status in medieval times.[2] This "diachronic gap" — but also the still incomplete description of the grammars and lexicons of ECL — almost inevitably led to multiple speculations on the genetic background of this language group. Their basic characteristics are the interpretation of areal conditions as being genetic, i.e., the assumption that all those languages of Caucasia which are not Indo-European, Turkic, or Semitic belong to the same language phylum, which is then divided into the three (or four) branches South, West, East (and North?) Caucasian. These speculations had become necessary also because they sometimes served as the basis for a broader comparison, i.e., the affiliation of the Caucasian languages with others outside the area (e.g. Basque, Hurro-Urartian, Hatti, etc.). This romantic (and, as it turns out, "nostratric") view of the monogenesis of all Caucasian languages was supported by some typological similarities, both in phonetics and morphosyntax. Though it is an obvious fact that typological similarity cannot serve as the basis for a genetic hypothesis, but may only secondarily induce speculations on it (cf. Schmidt 1962:6), it is often used to reinterpret a linguistic system which does not fit into the general typological frame of the Caucasian languages, simply by claiming *e silentio* that some features have been lost and must only be "reconstructed".

It must, however, be stressed that up to now we do not have any consistent proof that the Caucasian languages have a common genetic background. There are only a few lexical items which seem to have "cognates" in all three subgroups (cf. the list in Klimov 1969:55[3]), but these equations do not exhibit a system of regular sound laws, the only valid proof of genetic relationship we have up to now.[4] What has been claimed so far, is also true for the East Caucasian language group, though to a less problematic extent. Summarizing the genetic structures that have been found to be of relative certainty presents the following picture: the ECL have first to be defined negatively, i.e., they consist of all those languages, which are neither South nor West Caucasian. Five subgroups can be posited, with a greater or lesser high degree of probable genetic affiliation:[5]

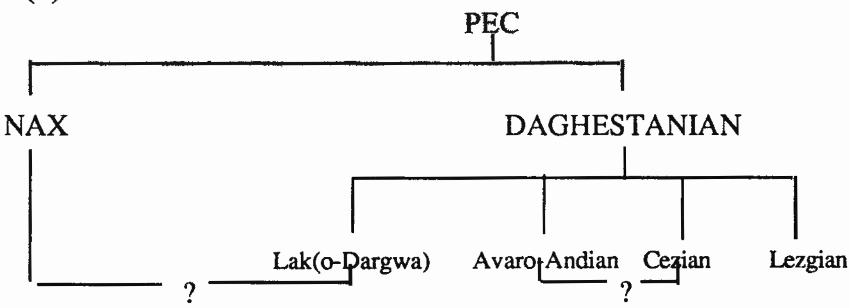
1. Nax languages: ((Čečen, Inguš), Bac));
2. Avaro-Andian languages: ((Avar (Andi, Botlix, Ĝodoberi, Karata, Axwax, Bagwalal, Tindi, Čamalal));

3. Cezian languages: ((Cez, Hinux, Xunza) (Xwarši, Bežita));
4. Lako-Dargwa languages: (Lak (Dargwa dialects));
5. Lezgian languages: (((Arči), (Udi), (Xinalug)) (((Lezgi (Aĝul, Tabasaran)) (Rutul, Caxur) (Kryz, Budux))).

The genetic relations within these subgroups are far beyond from being completely elucidated, but the following remarks may be made: The relationships *within* groups I to III cannot be doubted, although the genetic diversification in group II in particular (Andian languages) still remains unclear. Group V is well established for the so-called "Central Lezgian languages," which consist of the three branches of the Samur languages (Lezgi, Aĝul, and Tabasaran), the Western Lezgian languages (Rutul and Caxur), the Southern branch (Kryz and Budux).[6] Clearly of Lezgian origin also are Udi and Arči (Arči characterized by strong interference from Avar and Lak, which often led to its interpretation as "Avar"), whereas the status of Xinalug remains problematic, despite a relative high degree of probability regarding its Lezgian status. The Lak-Dargwa group, finally, is much more hypothetical than the others, a comparative study on these two languages (or, with respect to Dargwa, dialect groups) still is needed (a first attempt can be seen in Musaev 1968, though his data go beyond those of Lako-Dargwa).

An interpretation of the relationship *between* the above mentioned groups seems to be much more complicated. There is a close affinity between group II and III, often leading to the label "Avaro-Andi-Cezian" languages. But it still remains doubtful whether we should view the Cez group as a daughter or a sister of group II. Group V seems to have its closest genetic contacts with group II, whereas the status of Lako-Dargwa remains somewhat uncertain (also because of the doubtful internal structure of this group). Earlier assumptions which held that Lak had a closer genetic affinity to group I (the Nax languages)[7] later could be substantiated by a considerable number of lexical (and perhaps morphosyntactic) isoglosses proper to these languages only.[8] However, the reconstruction of the phonetics of Proto-Lak(o-Dargwa) hints, as far as we can see, at a common background with group II/III and V. The problem is directly connected to the question whether the Nax languages really are ECL. Although earlier proposals to treat them separately as a fourth Caucasian group can be rejected today, their status within the ECL remains in the dark. According to our present knowledge, it seems appropriate to regard them as a separate branch of PEC, opposing them to the proto-language underlying all other groups, often labeled "Daghestanian" or "PEC proper."[9] To conclude, a tentative (and highly abstract) *Stammbaum* may be presented as follows:

(1)



This *Stammbaum* suggests that all ECL have a common genetic background, called Proto-East Caucasian (PEC). Though the PEC-hypothesis seems to have much more probability than that of a common Caucasian proto-language, diachronic research still is far from being able to propose a coherently reconstructed system of the assumed PEC phonetics. The enumeration of valid sound laws based on correspondences between whole lexical items is the only way to prove a PEC background. If lexical items are segmented in order to make them "comparable," every step of segmentation should be explained by means of word formation and derivation. This restrictive method, the application of which has been so successful in Indo-European linguistics, is the only guarantee for avoiding root etymologies, fatal for any reconstruction project. Though the diachronic gap of the ECL complicates the situation considerably, there seems to be no reason to abandon the basic principles of sound laws and analogy (cf. Osthoff/Brugmann 1878:XIII: “Aller Lautwandel, soweit er mechanisch vor sich geht, vollzieht sich nach ausnahmslosen Gesetzen”). Yet we have to ask whether there are supplementary means which could help to fill this diachronic gap in order to better trace back the history of sound change in the ECL, leading to a comparative grammar of the East Caucasian languages.[10]

2. Functional Diachrony. I propose the term “Functional Diachrony” (FD) to refer to the evaluation of diachronic processes, which are joint to the functional categories of linguistic paradigms. By this is meant that parallel to the history of sound systems (sound laws and analogy)[11] we can assume a history of the functional properties of a given paradigm. Thus FD can be considered as the complimentary aspect of a formal diachrony, both of them characterizing the history of a linguistic sign, cf.

(2)

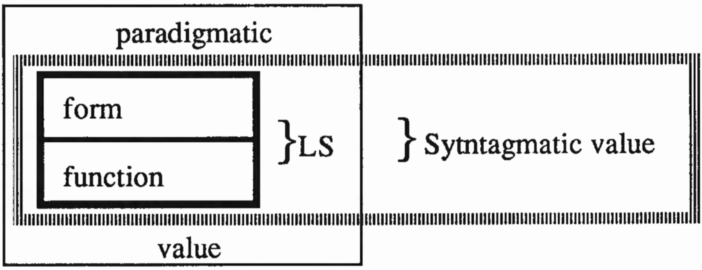
$$\Sigma_n \left(\frac{\textit{Form}}{\textit{Function}}\right. \rightarrow \Sigma'_n \left(\frac{\textit{Form}'}{\textit{Function}'}\right) \rightarrow \left(\frac{\textit{Formal}}{\textit{Functional}}\ \textit{Diachrony}\right)$$

$$t_0 \longrightarrow t_1$$

Σ_n = linguistic sign (n); Σ'_n = variant in time

The dialectic relation between form and function is not always equal when treating a linguistic sign diachronically. Especially when dealing with the shape and structure of an assumed sound system, its emic, i.e., functional, value may partly be neglected in the first analysis. However, when a linguistic sign is labelled with a grammatical value, i.e., when it is treated as a morpheme, then the functional component is of high importance. In this case, we have not only to look at the relation between form and function within the linguistic sign, but also at its place in an assumed paradigm together with its syntagmatic realization. Thus we have a threefold[12] organization of a given morpheme, which should be adequately treated in a gestalt-oriented reconstruction of linguistic signs, cf.

(3)



When reconstructing a linguistic sign, it is obvious that the primary basis is normally formed by its formal aspect. Languages with a long and well-documented, written tradition, as, for instance, the Indo-European languages, offer the possibility of drawing conclusions on the functional value of a reconstructed linguistic sign (LS) by its formal shape and its place in a specific paradigm.[13] Still, it is important to note that in the sense of FD we also have to check whether the reconstructed item fits into the more general frame of an assumed functional paradigm.

A Functional Diachrony does not neglect the findings within the realm of formal reconstruction. However, it tries to add assumptions on the functional characteristics of a paradigm which is thought to be the emic frame of a reconstructed item, together with efforts to explain diachronic processes within functional systems and subsystems according to proposals made by system theory and cybernetics. The explanation should be powerful enough to parse these processes in both directions (top-down and bottom-up) and should serve in all language domains (phonology, morphosyntax, semantics, pragmatics). Another important factor which dominates FD is the assumption that the distribution of functional values in a paradigm or between different paradigmatic subsystems probably is not equal, but structured according to the feature of markedness. The interaction between the prototypical kernel of a dynamic, self-regulating, partly open system (as may be proposed for language systems) and its periphery (characterized by vague boundaries) along the parameters "unmarked" (kernel) and "marked" (periphery) is of major importance for explaining a functional shift leading to the restructuring of a paradigm. This shift may take place (dialectally) in quality as well as in quantity, in expression, it may change the functional values of a paradigm, increase or reduce it, or — interacting with other (sub)systems — may lead to a syncretism with those. This final point is of importance especially when a paradigmatic change is induced by the restructuring of syntactic properties, e.g., by a shift from an ergative to a nominative type, which may affect the given techniques of agreement.

All these changes have to be described according to the more general principles of system change, as they have been first described in thermodynamics. Adapting the formulations of the two main postulates of thermodynamics,[14] we can claim that

> **1.** No functional value of a linguistic paradigm serving to maintain human communication can be lost in a diachronic process. It may, however, be substituted by nonverbal categories.[15]
>
> **2.** Every language system or subsystem diachronically tends to maximal "disorder," i.e., to a maximal state of entropy. System entropy is the decisive diachronic parameter for formal as well as for functional categories.[16] Its degree can be reduced by external, disturbing features (such as, e.g., the introduction of a script, historical events, etc.), stabilizing a given state of the system. Regarding the interaction of different language subsystems (as, for instance, agreement and case marking), FD imperatively has to look for isotropic and anisotropic processes, i.e., it has to ask whether the change of one subsystem takes place coherently with that of another one, functionally joined to it. Finally, FD should bear in mind that (in rare instances) a process may reintroduce older, previously lost categories (retropic processes) because of inherent structures of the system.

It must be added that the assumption of a functional change should follow the general principles of probability. This parameter not only includes typological generalizations on language universals, but also our knowledge of changes within the social background of a given language community. This is important, especially regarding those analogous processes which are not promoted by internal, systematic factors (i.e., by entropy), but by changes in actual world knowledge.[17]

Bearing these principles in mind, a Functional Diachrony may help to circumnavigate the diachronic gap in a given language group. Always trying to be consistent to already discovered phonetic processes, it can ask whether a supposition about the history of a given morpheme matches its functional properties and the structure of the paradigm with which it is associated. The basic task of FD is then to explain the single steps in the functional development of a system in order to draw conclusions on the (mainly morphosyntactic) structures of the proto-system.

3. The problem of East Caucasian class markers

3.1 General remarks.[18] The noun-class marking strategy is one of the basic morphosyntactic features of nearly all East Caucasian languages. With the exception of Udi, Lezgi, Aĝul, and Southern Tabasaran, all ECL have a more or less differentiated system of classifying nouns with the help of class markers (CM), cf. as an example:

(4)[19] Lak: bu-t:a-l b-a-w-x̂:u-nu b-ur ču-ø
father-SA-ERG III-buy-AOR III-AUX horse-ABS
"the father bought a horse"

barzu-lt-ø-wa bak′qaqu-lt-ø bu̲x̂t:ulnu na̲k′-s:a s:aw-n-iy hawa-b-ull-ay
eagle-PL-ABS-and vulture-PL-ABS high blue-ATTR sky-SA-SUPER/LOC air-III-make/DUR-PRES
"Eagles and vultures cruise high in the blue sky."

These examples show that the absolutive-marked nouns in both phrases are referred to by means of the class marker *b-*. In the first (transitive) sentence it encodes the third class singular, in the second (intransitive) case — the third class plural. As can be seen, the class marking technique in ECL possesses paradigmatic (or semantic) as well as syntactic properties. I do not want to discuss these techniques in detail here,[20] but will briefly enumerate the basic characteristics.

1. The basic technique of noun classification (NC) in ECL is covert, i.e., nouns controlling class agreement normally are not marked by CM.

2. Class marking involves attributive as well as verbal agreement. The latter is canonically restricted to the first argument of a predication, i.e., to absolutively nominals encoded in the absolutive. Single language processes may lead to secondary (?) marking techniques (for instance, datives, possessives etc.).

3. Paradigmatically, the number of classes goes from two (Northern Tabasaran) up to ten (Čečen dialects)[21]. The basic characteristics of plural CM is the use of the singular inventory.

A typical paradigm of class morphemes can be seen for instance in Arči (marginal Lezgian language). The suffixed series with verbs draws following scheme:

(5) Arči:

	SG	PL
I	-w	-b
II	-r	
III	-b	
IV	-t(:)	-t(:)

(5) shows the heteromorphic structure in the singular, which is encoded by means of distinct morphemes, whereas the plural copies the singular morphemes together with the typical reduction of differentiation, as can be observed in many plural paradigms.[22] Morphologically, the plural morphemes refer to the singular classes III and IV, whereas the singular classes I and II remain isolated and marked, cf.:[23]

(6)

class	SG	PL
I	a	c
II	b	c
III	c	d
IV	d	d

Before turning to this type of subcategorization, we have to take a brief look at the syntactic background of the class-marking technique in ECL. I have discussed this problem in Schulze 1988, and thus will constrain myself to the most relevant facts.

The syntax of ECL class marking technique can best be described in terms of an extended prototype theory. By this is meant that — adapting the concept of semantic prototypes as developed by E. Rosch[24] and others — synchronic structures that genetically related languages have in common may reflect an older, structural module which represented the conceptional or prototype kernel of this category. In this sense, a diachronic prototype would be the common denominator of all single language structures in question which represented a probable cognitive reality in the proto-language. Thus, the prototype of a morphosyntactic structure would give us information on the basic prelinguistic schemes referred to paradigmatically and syntagmatically by the speakers of the proto-language, whereas the vague periphery (i.e., the contextual usage of the prototype) remains unexamined by FD in the first analysis. A diachronic prototype describes the central, nuclear features underlying a given construction, bearing in mind that this description is an abstract, metalinguistic, formulation of a probable linguistic reality.[25]

In this sense, the prototypical background of ECL class marking, which can be deduced from the above given three characteristics (among others, naturally), may be formulated as follows: The technique of class marking had its origin in the topicalization of the first argument in a proposition by means of anaphoric pronouns.[26] The prototypical structure of the nuclear proposition would be:

(7) $\text{*NOUN}_{\text{i/ABS/top}}$ — $\text{DEICTIC PRONOUN}_{\text{i}}\text{-AUX}$ — $\text{VERB}_{\text{/part}}$[27]

Thus, the impersonal verb, marked by a temporal/aspectual form, was first linked to its argument by means of an auxiliary, traces of which can still be found in many ECL. A typical intransitive phrase like:

(8) Čečen: stag- w-öl-u "the man is laughing"
man-ABS I-laugh-PRES

could be interpreted (in the sense of FD!) as "*the man, he is laughing." This construction can be regarded as the once marked variant of a type "*the man is laughing," which underwent a shift of markedness.[28] The shift then led to the interpretation of verbal forms as distant attributes of the first nominal (or pronominal) argument; it was joined to it as other attributes by means of agreement. The old deictic elements thus became "class markers." This process automatically affected the ergative construction, as its decisive characteristics, the adding of a control-marked nominal in an ergative case did not form part of the central predication (with labile verbs), but was added to it in object function (in terms of Dik 1989!). The transitive prototype would show following picture:[29]

(9) $\text{*NOUN[+control]}_{\text{j/ERG}}$ ($\text{NOUN}_{\text{i/ABS/top}}$ - $\text{DEICTIC PRONOUN}_{\text{i}}$ - AUX - $\text{VERB}_{\text{(trans)/part}}$)

A transitive proposition like (4) could be interpreted as "*The father causing that the horse, it was involved in the action of buying," a highly marked and analytic construction, which necessarily tended towards agglutination. The resulting mechanism was the "canonical" agreement of the absolutive marked agent/patient with the verb by means of then class markers.

The use of anaphoric pronouns, which referred to the first argument of a proposition, automatically led to the semantic subcategorization of the category "noun." It is because we have a paradigm of noun classes together with their specific morphemes, that we cannot assume a single multifunctional deictic morpheme, but rather a set of demonstratives which came into use in this construction. Thus, the number of classes was generated according to the number of pronominal elements in use. This explains why we do not have a numerically open category of noun classes as in typical overt systems, e.g., in Bantu languages. Prototypically, the subcategorization generated four distinctive noun classes in PEC.[30] Their class morphemes copied semantic features of nouns parallel to the anaphoric procedures of deictic pronouns. We can assume that the underlying procedure of semantic categorization took place according to social and cognitive parameters. In the sense of FD it is important to note that though we can observe a general tendency in ECL towards a more or less distinct four class system, a consequent parallelization of class allocations in the different languages seems to be impossible. There surely are certain common features, cf. the trend to encode the features [+masc; +human] by class I,[31] but in general hardly any allocational noun type has its exact correspondence in more than three languages. This is why we must presume that the procedures of semantic allocation were essentially based on social convention and knowledge of the world. It seems that the distinctive semantic features represented a *n*-tuple of vague associations which could easily be changed if their original functional value was no longer relevant or if conventions also changed.

Reviewing lexical lists in the individual languages together with an analysis of the strategies for allocating nouns to the different classes hints at a basic feature [±social relevance]. By this is meant that nouns in PEC probably were allocated according to whether their denotata were relevant for the maintenance of social structures and communication. This inferential feature was supported by secondary ones, which scaled it according to a specific hierarchy. As can be shown from other paradigmatic factors in ECL,[32] this hierarchy was dominated by the "degree of possible intervention" in a process (DPI, *"Eingreifgrad")*, i.e., by the question whether the denotatum of a noun was thought to be able to cause a process or to intervene in one. This feature corresponds more or less to that of [±control], though it cannot be interpreted binarily as the latter can. Within this hierarchy, class I with the highest DPI represents the most marked category, which allowed hardly any or only few deviations from the prototypical semantics, whereas classes III and IV are relatively unmarked, showing a high rate of semantic diversification. The following table sums up the DPI hierarchy:

(10) HIGH <---------- DPI <--------------- LOW
CLASS: I <<<<< II <<<<< III <<<<<<<<<<<<<IV
marked < < < < < < < < < < < < unmarked

The changing DPI corresponds to the fact that class I has the highest functional load. Its strong emic character hardly allows any allotypes, whereas class III and IV — on the contrary — emicly are less relevant. This matches the high rate of allotypes found in these classes. The distribution of denotational and inferential features in the prototype of PEC noun classes can be listed as follows:

(11)

DENOTATIVE FEATURES	**CLASS**	**INFERENTIAL FEATURES**
[+masc; +hum]	**I**	[+control]
[-masc; +hum] [-hum; ± anim]	**II**	[+control$_{1-\emptyset}$] [-control; +relevant {+fem; +hum}]
[αmasc; +hum] [-hum; ±anim]	**III**	[-control] [-control; +socially relevant {+hum}]
[-hum; ±anim]	**IV**	[-control; -socially relevant {+hum}]

The left column corresponds to the traditional assumptions on PEC classes, which are characterized by a binary interpretation (cf. Deeters 1963:46).[33] It neglects however the fact that especially in the more unmarked classes II-IV many "atypical" allocations can be found. This is especially true for class II, which comprises many denotata clearly belonging to the working sphere of women (in the Eastern Caucasus).[34] Also the use of class III with lexemes for "child" is often stated without comment. Obviously we have to think of an important cut-off point according to features like [±able to speak], [±young], [±initiation], [±unmarried], etc., depending on the social structure of the language community.[35]

The problem of allocating the features [±relevant] and [±control] plays a considerable role in describing the singular-plural proportions in the class system. In a prototypical sense, the marking of plural agreement would not be economical, as the once topicalized nouns already had a distinct plural morphology. This assumption matches the fact that — as mentioned before — the plural morphemes almost exclusively belong to the singular stock. This indicates that the plural CM originally did not refer to plurality, but to specific inferential features which comprise both singular and plural. The tendency to reduce the functional categories openly marked in the plural (underdifferentiation of the plural) reminds us of the use of the most unmarked CM in the plural. Only these contain the necessary minimum of distinctive features (and of DPI, see above), so that they can be introduced in a plural paradigm which eliminates distinctive features for discrete objects of the world. The concept "plural" in PEC obviously had a collective interpretation, which can be deduced from many indications within the paradigm of stem augments and from restrictions on antipassive formations.[36] Thus, in order to encode plural agreement, PEC made use of the least distinctive, nearly collective classes of the singular, i.e., classes III and IV. In class I and II, the distinctive (and discrete) feature [+masc] has been deleted, generalizing the feature [+hum]. The plural class I/II then was encoded by means of the singular class III, because this class did **not** refer to [±masc], but positively referred to [+anim]. Collectives of class III and IV, however, found a less animate interpretation, emphasizing the economical ("non-animate") relevance of their members (e.g., flocks of sheep are interpreted to be "less animate" than a single, discrete sheep).[37] Thus, the plural shows a mixture between denotational and inferential features, resulting in the following picture:

(12)

DENOTATIVE FEATURES	**CLASS**	**INFERENTIAL FEATURES**
[±masc; +hum]	I/II	[+control]
[-masc; ±hum]	III/IV	[-control; ±relevant]

3.2 The process of petrification. Having enumerated the basic features of class marking strategy in ECL together with its underlying prototypical aspects, I now want to turn to the question of why and how CM can petrify. This problem is of considerable importance regarding proposals to reconstruct PEC lexical items. When looking at these proposals, we can observe three basic assumptions:

1. All ECL have or have had a system of morphological noun classification.

2. Both in languages that have maintained this strategy and that have lost it we are able to trace former CM now incorporated in either nouns, pronouns, verbs, or adjectives.

3. A considerable number of inflectional morphemes, such as plural and some case markers represent former CM.[38]

If these assumptions turn out to be true, in fact, the class marking strategy would be the most important feature of PEC morphology, dominating nearly all other paradigmatic categories. However, a closer look at both the scientific interest of the investigators in question and the linguistic evidence show that these proposals are much too dominated by *ad hoc* presumptions. Still, point 1 probably was true for the proto-languages of the different ECL groups and for PEC. This can be deduced from the following facts:

1. The four languages Udi, Aĝul, Lezgi, and Southern Tabasaran, that do not show a system of noun classification, all belong to the Lezgi subgroup.

2. The morphosyntax of Southern Tabasaran is contrasted to that of the northern dialectal variant, which has a (however, quite decayed) system of NC.[39] The morphosyntax of Southern Tabasaran in general is younger than that of Northern Tabasaran.

3. Lezgi, Aĝul, and Tabasaran have both a close genetic and areal affinity. We may claim that the fact that they do not have a NC system is due to a *Sprachbund* phenomenon. However, it must be stressed that, in spite of this common feature, there is a remarkable difference between Lezgi and Aĝul on the one hand, and Tabasaran on the other. The first two languages have lost all means of encoding agreement between the verb and its arguments, whereas Tabasaran has developed a quite complex, partly polypersonal, system of personal agreement.[40] Thus, it might be doubtful whether we have to deal with a genetically based "negative innovation" in the proto-language of this group,[41] or with two different types of functional processes.

4. One of the most marginal (and in many respects archaic) languages of the Lezgi group, i.e., Arči has a complex system of NC, quite close to the above mentioned PEC prototype.

5. The data of the Western and Southern Lezgian languages together with those from Arči and Xinalug allow us to morphologically reconstruct the system of NC in Proto-Lezgian.
6. The loss of NC in the quite marginal Udi should be explained by single language processes, perhaps due to the restructuring

of Udi syntax (nominativization) and to language contacts with Azeri and Armenian.[42]

Still, it has to be stressed that the assumption of a class marking technique in PEC is mainly based on typological evidences. The proof that we have to deal with a common, genetically related system in all ECL can only come from a formal comparison of the different morphemes is use. In fact, it seems that the nuclear system of CM in PEC made use of the following morphemes:

(13)

	SG	PL
I	*w	*b
II	*y/*r[43]	*b
III	*b	*d
IV	*d	*d

Perhaps it is not just a coincidence that the first two classes are marked by sonants, whereas III and IV have voiced stops. This fact corresponds to the above described distinction between the marked and unmarked elements in the system (cf. (6) above). Peripheral realizations of noun classes in PEC, which may have had considerable impact on the single language systems surely have existed, but unfortunately no research on this problem has been done up to now, so that I have to concentrate on the nuclear morphemes. Thus, whenever an element of a given paradigm is considered a petrified class marker, the first condition should be that its formal aspect corresponds to that of one of the reconstructed morphemes according to the sound laws of the language. This restriction already eliminates a considerable number of cases in which researchers have claimed PCM.[44] The general tendency that can be observed here is to interpret an element as PCM regardless of its phonetic history. Nearly every (reconstructed or existent) phoneme has already been referred to in order to support a PCM interpretation. Even without incorporating functional aspects, it can easily be claimed that such an interpretation has to be rejected at least in the following instances:

1. if the phonetic shape of an assumed PCM is absolutely incompatible with the reconstructed morphemes (cf. above) according to general ECL sound laws. Thus, e.g., a uvular or velar (both stops and fricatives) could never have been a PCM.[45]

2. If the phonetic shape of an element in question meets that of a reconstructed CM or its phonetic variant, but does not show a parallel phonetic environment. E.g., in some cases, CM III **b-* can become *-m-*. But this happens only in medial position (as in Lak), more concretely: **b > m/C_V*. This change never occurs initially, so that all efforts to explain a root structure *mVC-* (well known in ECL) by *CM(III)-VC-* should be abandoned.

It has to be stressed, however, that especially in the Western Lezgian languages Rutul and Caxur there are very complex types of combinatory sound change in which CM are involved. The results are sometimes completely "disguised" structures which may also have existed in other ECL. A typical example is the Caxur paradigm of *eč′es* "to enter":

(14) Caxur:		SG	PL
	I	eč′es (< *e-ø-č′es)	e-b-č′imĕ
	II	ēč′es (< *e-y-č′es)	e-b-č′imĕ
	III	ōč′es (< *e-b-č′es)	e-d-č′imĕ
	IV	eč′es (< *e-ø-č′es)	e-d-č′imĕ

It is remarkable (and not yet explained) how the plural CM *-b-* and *-d-* are maintained in nearly the same position in which they are assimilated (or lost) in the singular. The process of assimilation as it is documented in the singular classes II and III is of considerable importance not only in Caxur and Rutul, but also in other ECL (Nax and Cezian languages, the Balxar dialect of Lak, Xinalug). Fortunately enough, we often can associate such processes with a given system of noun classification. But we have to assume parallel processes also in paradigms which today are no longer involved in this system. In fact, it might be much easier and more attractive to declare, e.g., Udi *be̲ĝ* "sun" to be representing a structure *CM(III)-*e̲ĝ* than to discuss the concurrent Lak pair *zix̂:in /x̂:ix̂:in* "to rob" in terms of CM.[46]

Lexicalization of once paradigmatically relevant elements is a common trend in human language. Thus, assumption (2), in positing the existence of petrified class markers seems to have a certain probability. However, when scanning the relevant literature on PCM in ECL, it becomes obvious that the label PCM often serves as a "cure-all" remedy in those instances when the formal comparison of two or more lexical items will lead to a reconstructed form. The general trend is to reduce the size of lexemes by analyzing them as a composition of class marker and stem. This procedure often leads to root etymologies, which — as has been said above — can hardly serve as diachronic evidence. This way of analyzing PCM in ECL can be generalized as

(15) $C_xVC\ (L_1) = C_yVC'\ (L_2) \Rightarrow -VC = -VC'\ \&\ C_x \neq C_y \Rightarrow C_x/C_y\ \varepsilon\ \{PCM\}$,

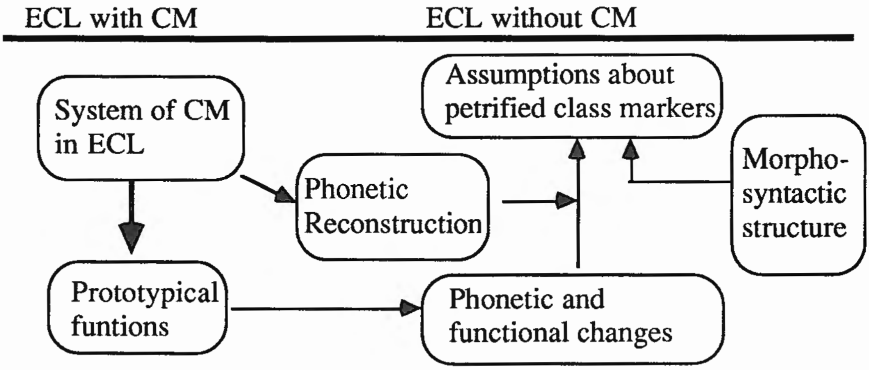
which reads: If two lexemes from two different languages are compared, and if a part of them matches given sound laws (C = C′), then the rest probably belongs to the set of class markers. A typical example of such a procedure is given by Talibov 1980:306f.: looking at the correspondences for "two" in the Lezgian languages he observes a velar or uvular component in all entries but the Udi, cf.:[47]

(16)

Lezgi	q′°ed	Arči	q′°e-CM
Tabasaran	q′u(-CM)	Kryz	q′a-CM
Aĝul	q′ud	Budux	q′a-CM
Rutul	q′°a-CM	Xinalug	k′u
Caxur	q′o	Udi	p′a

Because there are some (but very uncertain) hints at a possible loss of final uvular stops in Udi, Talibov proposes that Udi *p'a̲* emerged from **p'aq'*°, which then is confronted with the data of the other languages. He concludes, in line with assumption (15) that the basic form in Udi was a class marked **b'aq' ˁ̊(a)*. It can be easily shown, however, that Udi *p'a̲* is a regular reflex of proto-Lezgi **q' e̊-*, according to the Udi sound law **q' ° > p'*.[48] Moreover, the examples show that the class-marking strategy with numerals is based on suffixation and not on prefixation (cf. the dentals in Lezgi **q' e̊d* and Aĝul *q'u̲d*, which obviously are PCM).[49] Still, the assumption of petrified class markers can be helpful for determining the history of lexical and morphological items in the single ECL. But we have to ask which conditions must be met so that a given element can be regarded as a PCM. In other words, we have to draw conclusions from both the formal reconstruction of CM and their embedding into the functional frame of the assumed PEC syntax (i.e., its diachronic prototype). Secondly, if these conditions are met, we must explain why and how the process of petrification came about. (17) gives a very rough picture of the procedures necessary to determine the possible status of an element as a PCM:

(17)


ECL with CM
ECL without CM
Assumptions about petrified class markers
System of CM in ECL
Morpho-syntactic structure
Phonetic Reconstruction
Prototypical funtions
Phonetic and functional changes

Thus, we have to start with the paradigmatic systems in those languages that still have a system of NC. With the help of sound laws already established for the ECL, it seems possible to reconstruct the prototypical nucleus of the system of NC in PEC. This system was proposed in (13). This formal approach has to be combined with assumptions abut the prototypical functions of class markers in PEC, as is shown in (8) and (9). Now, both formal and functional changes from PEC to the individual language groups and from there to the single ECL must be traced and described before the resulting propositions for PCM can be equated with the data from languages without PCM and their internal morphosyntactic structure. Naturally, the same procedure holds for assumptions on the lexicalization of CM in class-marking languages, also. The possibilities of finding PCM seem even more promising in these instances, because we can immediately draw a parallel between the element in question and its internal paradigmatic counterpart. This formal advantage, however, is somewhat reduced because of the fact that the functional probability of petrification in a vital system of NC is less high and should be explained separately.

The prototypic scheme of NC in PEC has to be described according to its paradigmatic and syntagmatic aspects. Paradigmatically, we can assume a four-class system, as has been claimed above. The problem concerning PCM is, then, which class markers within this system have the highest probability of petrification and which have the lowest. In order to determine this degree of probability, it is necessary to take a closer look at what is called "lexicalization." Normally, this phenomenon is treated as a part of diachronic word formation and describes the processes and results of incorporating once-independent morphemes into a another lexeme. This process coincides with the morphological demotivation of an element. Thus, one often has recourse to lexicalization when observing tendencies for change from a complex, analyzable group of morphemes to a single lexical unit, the elements of which can no longer be deduced from its semantics.[50] Lexicalization can take place both on a purely lexical and on a syntactic level. The lexical level is of less relevance for PCM in ECL, though there are some interesting findings here, too. Thus, Udi *x̂unči* "sister" can be paralleled with Tabasaran *či*, Rutul *riši*, Caxur *yiši*, Kryz *šiděr*, Budux *šider*, Arči *došdur*, and Xinalug *rěc*^{h} (all with the same semantics). It can easily be shown that we have to reconstruct a proto-Lezgi form **rěc(ě)*, which has good correspondences in nearly all other ECL.[51] I do not wish to enter into the discussion of the individual lexemes, but it is important to note that the Kryz and Budux forms represent pluralia tantum and have lost the prefix **rĕ-* as in Tabasaran. **rĕcĕ* together with proto-Lezgi **wĕcĕ* 'brother' forms a morphosemantic paradigm which has its parallel in the PEC pair **weš(:)ĕ* 'son' and **r/yĕš(:)ĕ* 'daughter.' All four items have PCM, prefixed to the stems **cĕ* and **š(:)ĕ:*

(18)

I	*wĕčĕ	*wĕš(:)ĕ
II	*r/yĕcĕ	*r/yĕš(:)ĕ

The semantics of **-cĕ* probably was something like "the one belonging to the same family," whereas **-š(:)ĕ* perhaps meant "born." Thus **r/yĕcĕ* referred to 'the female member of the family' > 'sister.' Udi has lost initial **r(V)-* regularly. Thus, we have to assume an intermediate state **rĕcĕ* > **ci*, whereas **wĕcĕ* still today is reflected by Udi **wiči* 'brother.' Udi *x̂unči* can now be explained by the restoration of the old PEC structure, i.e., by the use of a morpheme denoting [+fem]. This was done with the help of Udi *x̂uni* (< **Λĕn/dĕ)* "woman.'[52] The lexicalization of *x̂uni* and **či* obviously copied the underlying paradigmatic structure of NC by means of lexical items.[53] In terms of system theory, Udi gives us a clear example of what can be called a "retropic" process, which, however, could not stop the trend of loosing NC in Udi.

Yet, lexicalization of CM on a syntactic level seems to be the most dominant factor with respect to the functional processes of CM petrification in ECL. In the sense of FD and applying the assumptions to the formal conditions of CM both in ECL and PEC we have to ask how and why CM were incorporated in the lexemes of a given language. Let us look at an example from the Lezgian languages. Udi *buq'-sun* 'to love' has the following parallels in the Lezgian languages: Lezgi *k'an x̂un*, Tabasaran *k:un x̂uz*, Aĝul *k:an x̂as*, Rutul *yigĕn*, Caxur *yik:an-as*, Kryz *ik:-äǯ*, Budux *yika*, Arči *λ:ʹan*, Xinalug *yuk:ı̊*. Though word formation rules have led to quite different results in the individual languages,[54] it can easily be shown that the underlying root in proto-Lezgian was **λ:'ĕ-*.[55] But what are the prefixes found in the non-central languages and Arči? Whether they are CM or not can best be observed in those languages that still have NC. For instance, the Caxur paradigm of *yik:anas* shows that there are tense/aspect forms without the prefix: CM appear in its place, cf. the negated forms. In the present tense, we have a morpheme *yi-*, which could be equated with the prefix, but the latter does not take part in the CM paradigm (i.e., we do not have infinitives like **ı-k:anas* and *wu-k:anas);* cf.:

(19) Caxur:

	'not to love'	'loving'
I	de-Ø-k:anas	ı-k:an
II	de-y-k:anas	yi-k:an
III	de-w-k:anas	wu-k:an
IV	de-Ø-k:anas	ı-k:an

It follows that, for structural reasons, the prefix in *yi-k:anas* cannot be a class marker. It was shown recently that we are instead dealing with an aspect marker, which, for instance, appears in prohibitives *(mi-y(i)-k:an* 'do not (start to) love').[56] Parallel claims can be made for Rutul, and there are some hints at an aspectual origin of the prefixes in Kryz, Budux, and Xinalug, also. Thus, the palatal prefixes must be dropped as candidates for PCM. What about the Udi form *buq'sun?* As Udi no longer has NC, we might think of a PCM status for *bu-*. This does not mean that *bu-* could not have something else, but two facts lead to the assumption that it was a PCM. First, Arči *λ:'an* as well as the Central Lezgian data show that something must have been added to Udi *buq'sun* (< **bu-λ:'(ĕ)-)*. As we do not have clear evidence for a preverbal or aspectual morpheme **bV-* in any proto-Lezgian language, its status as PCM becomes more likely (also because of the fact that its formal shape finds its parallel in the system of reconstructed CM; cf. above). The underlying structure *_CM-λ:'_ - is still preserved in Avar, which has _CM-óλ:'-ize_ 'to love.' This form is regularly inflected according to NC; cf.:

(20) Avar:[57]

I	dun wac: w-óλ:'-arab	"I love the brother"
II	dun č:'užu y-óλ:'-arab	"I love the woman"
III	dun ču b-óλ:'-arab	"I love the horse"

Interestingly enough, the citation form for the Avar lexeme normally is _bóλ:'ize_ (infinitive), thus the class III marked verb is used in a "neutral" sense.[58] We may thus infer that at least in Avar, class III is treated as the least marked, which means that its functional and semantic properties belong much more to the periphery of the NC system than those of class I and II. This must not always have been true, as it can shown from parallel forms, for instance, in the Cezian languages. Here 'to love' is cited by Bokarev 1959:233 as _y-eta_ (Cez), _y-at'a_ (Xunza), and _y-at'al_ (Bežita), i.e., the citation forms are based on class II. [59]

With the help of this typological and perhaps genetic parallel, the assumption that Udi _buq'sun_ contains a petrified class marker has become highly probable. We should now ask how the process of petrification can be described in this instance and whether we can observe functional restrictions which would clarify it.

The first problem is: which categories of lexemes in ECL are liable to have retained older CM. It seems a likely claim that only those (functional) word classes can be marked by PCM that once were involved in the syntactic strategy of NC. The prototypical system of NC has been characterized as being "covert," i.e., nouns controlling their classes were not marked by the corresponding class markers themselves. We can draw two conclusions from this statement:

1. Prototypically, all word classes that can be marked by CM may show PCM.

2. When assuming PCM for an actual noun, the noun must not have been a noun when the class marking technique still involved this lexeme, i.e., {PCM}LEXEME[+NOUN][60] < *CM-LEXEME[-NOUN]. The lexemes in question must have belonged to word classes deduced from conclusion (1).

The underlying technique of chaining attributes (adjectives or pronouns) and verbs with a noun by means of CM leads to the conclusion that we should initially look for PCM in these categories (as has been done for Udi *buq'sun)*. The situation is, however, much more problematic and (accordingly) disputed regarding PCM with nouns. In nearly all grammatical works on ECL we can find relevant proposals, the most radical is perhaps the formulation of Charachidzé 1981:31: "La présence de ces marques fossiles dans de nombreux lexèmes nominaux conduit à une certitude: dans un état antérieur de la langue (s.c. Avar), les noms portaient la marque de leur appartenance à telle ou telle classe de l'univers conceptuel." It must be stressed, however, that this *"certitude"* by no means fits into the general scheme of prototypical noun classification and ergative construction in ECL, as has been described in 3.1. In order to approach this problem, we should first take a look at those languages which overtly use CM with nouns so that we can determine their categorial status and evaluate the possibilities of overt NC in PEC. I shall cite two examples only: First, an often quoted Avar paradigm, which shows an interaction between a noun and CM; cf.:

(21) Avar:

I	w-ac:	brother	w-as	son
II	y-ac:	sister	y-as	daughter
III	b-ac:	animal of the same generation	b-as	young one

As has been claimed above, this technique of labeling the concepts <CHILD> and <BROTHER and SISTER> is quite common in ECL and belongs to PEC word formation strategies. We now have to ask whether we can assign the marker [+NOUN] to the Avar stems *-ac:* and *-as,* as is often done *e silentio.* Obviously, it is the class marker that determines the semantic concretization of the concepts given above regarding the feature [±masc; ±hum], whereas the stems seem to be neutral. If we have to deal with an overt classificational type, there seems to be a conflict in the assumption that a noun controlling its class controls with the help of its class marker, but itself is semantically underspecified. We may think of a compositional type as in English *he-goat* vs. *she-goat,* but this normally differentiates sex in a lexeme which already is marked by sex.[61] The roots *-ac:* and *-as,* however, lack the distinct semantic feature [±masc; ±hum], which could induce a technique parallel to the English example. Thus, we should abandon this interpretation. A proposal consistent with the above-mentioned conclusions would be to interpret these roots as former adjectives denoting concepts like <THE SELF> and <THE BORN>. The use of CM, then, can easily be explained by the use of these lexemes as attributes of later deleted nouns, which — semantically marked — specified sex and covertly controlled their classes accordingly; thus :

(22)[62]

w-ac: < *w$_i$-ac:$_{ATTR}$ {NOUN$_i$[+masc;+hum]}
y-ac: < *y$_i$-ac:$_{ATTR}$ {NOUN$_i$[-masc;+hum]}
b-ac: < *b$_i$-ac:$_{ATTR}$ {NOUN$_i$[±masc;-hum]},[63] etc.

Interestingly enough, we have proof that these lexemes have been or still are treated as attributes in at least one ECL. The data come from Arči (a Lezgian language). There are 14 nominal lexemes in Arči that are encoded by means of CM. All of them denote either kinship terms or professions, among them the two pairs "brother/sister" and "son/daughter"; cf. the following table:

(23)

Class I [+masc; +hum]		Class II [-masc; +hum]	
ušt:u	brother	došt:ur	sister
eyt:u	father	eyt:ur	mother
šekert:u	cousin	šekert:ur	cousin
iškert:u	2nd cousin	iškert:ur	2nd cousin
eyšekert:u	3rd cousin	eyšekert:ur	3rd cousin
(w)išdu	boy	(d)išdur	girl
mac'at:u	groom	mac'at:ur	bride
nust:u	brother-in-law	nust:ur	sister-in-law
d'aut:u	neighbor	d'aut:ur	neighbor
elet:u	guest	elet:ur	guest
uΛdu	shepherd	uΛdur	shepherdess
c'ix̂dit:u	thief	c'ix̂dit:ur	thief
aršit:u	Arči man	aršit:ur	Arči woman
baqaršt:u	potter	baqaršt:ur	potter

All these nouns are characterized by the formant *-t:u-* (or its allomorph *-du-)* which is used as a stem augment for adjectives in attributive position. To this formant the CM are added: Class I *-w* > *-0,* Class II *-r*. It seems that the starting point for this paradigm was the above-mentioned pairs "brother/sister" and "boy/girl," which still show the PEC use of prefixed CM *(ušt:u < *w-uš(-t:u-w), došt:ur < *r-oš(-t:u-r), (w)išdu < *w-iš(-t:u-w), (d)išdur < *r-iš(-t:u-r))*. The forms *mac'at:u* and *mac'at:ur* copy the procedure to link CM-marked attributes with a later deleted noun on a more recent level. They simply belong to the adjective *mac'a* 'new,' thus denoting "the new (one in the family)." Much more work has to be done regarding the other lexical stems used in this paradigm. *Nust:u(r)* in particular shows that the technique could later be applied to nouns too, as long as they belonged to the realm of kinship terms. *Nus-* probably is a loan from an Indo-Iranian language, the source still being uncertain.[64] There seems to be another possibility when nouns appear class marked. This is the case in the very exceptional technique in Upper Andi of encoding the genitive: cf.:[65]

(24)

I	imu-w woc:i	"the man's brother"
II	imu-y yoc:i	"the man's sister"
III	imu-b k'oti	"the man's horse"
IV	imu-b qanǯar	"the man's dagger"
V	imu-r haq:'u	"the man's house"

The underlying form is the chaining of *determinans* and *determinatum* (in this order!) not by means of a genitive morpheme, but only by CM. This structure can also be found in Karata (Andian).[66] In Andi, it is confined to possessors belonging to Class I, whereas others are encoded by means of the genitive; cf. *ilu-λ:i woc:i* (mother-GEN brother) 'mother's brother.' The starting point for this type of genitive was probably the topicalization of (only human male) possessors at a time when CM still were deictic pronouns. Thus *imu-w woc:i* could be interpreted as *$\text{NOUN}_{i/\text{TOP}}$ DEIXIS_i NOUN (the $\text{father}_{\text{TOP}}$, (t)his brother). We may assume that in some rare cases the determinatum could be lost, as the semantics of the possessor already comprised most of its features. This process which would correspond to such English expressions as 'at the barber's,' 'at John's,' etc. might have led to the petrification of class markers with [+masc; +hum] nouns.

Before taking a look at some other nouns which often are analyzed as consisting of a noun and a class marker, we have to turn to the question of whether there are other syntactic conditions which can determine the fate of a CM as a PCM. As we have seen above (3.1), the prototypical technique of class marking was directly (and, because of its origin, automatically, joined to the prototype of the ergative construction. CM copy properties only of intransitive agents/patients and transitives agents (cf., again, (4)). If we claim that petrification results from the deletion of a controlling noun, we must conclude that:

> If we assume a verb to have lexicalized a CM, the syntactic and semantic conditions of the ergative construction must be met accordingly.

This conclusion is essential for any attempt to explain the process of petrification. Adapting the formulas given in (7) and (9), we can posit the following assumption:

(25)

$$\begin{aligned}&(\text{NOUN}_j\text{-ERG})\ \text{NOUN}_i\text{-ABS}\ \text{CM}_i\text{-VERB} >\\ &\qquad (\text{NOUNj-ERG})\ \{\text{NOUN}_i\text{-ABS}\}\ \{\text{CM}_i\}\text{VERB}\end{aligned}$$

Thus a petrified class marker in a verb should always — in its ancient function — refer to the absolutive marked noun then deleted, i.e., to the intransitive agent/patient or transitive patient. The process of deleting the noun of a proposition, however, is bound to several conditions:

1. **Frequency:** A noun may be deleted only when it frequently occurs with its predication, i.e., the more the construction is stereotypical, the more it is likely that parts of it may be dropped.

2. **Unmarkedness:** The nouns in question should have a very small semantic periphery. Their prototypical nucleus should cover most of their semantics. If a noun with an extended semantic periphery is deleted, we probably are dealing with a metaphoric construction.

3. **Information maintenance:** No process of deletion should affect the information structure of a predication (analogical to the postulate 1 in Section 2, above). Deletion can be understood as a process of reducing the size of an expression without reducing its quality. The listener should be able to infer the semantics of the deleted noun from his knowledge of the world, the semantic structure of the rest of the proposition, and assumptions about the probable semantics of the deleted category.

4. **Semantic vagueness:** The more a semantic concept is discrete, the less it can be deleted. This follows from (3), as discrete concepts bear much more unique information then others. Thus, names, deictically marked singular nouns, and pronouns can be deleted only with difficulty[67] (contrary to plurals and collectives with strong nuclear semantics).

5. Low DPI: Only nouns with a relatively low "degree of possible intervention" (cf. 3.1) should be considered as candidates for deletion. The more a noun controls its predication, the less likely it is to be deleted.

6. Semantic adequacy: Consistent with (3), the semantics of a given predication should already contain most of the inferential features of a deleted argument.[68] These features should meet with the prototypical nucleus of the verbal semantics.

These conditions represent the functional guideline for any procedure for discovering PCM. Any finding that forces us to assume a PCM but which is not consistent with them should be explained separately. We now have to ask which classes (together with their class markers) contain semantic features appropriate or inappropriate to these conditions. It has been claimed above that the system of noun classes was prototypically organized along the DPI or [±control] scale (cf. (10)). Class I probably was the most marked and most discrete category. The deletion of Class I nouns would lead to a massive loss of information, as predications based on them would rarely occur stereotypically. Nearly the same is true for Class II, whereas the parallel use of Class III and IV for singulars and plurals already indicates that they represent less discrete and much vaguer categories. Thus, we can establish a hierarchy within the system of NC with regard to the question of which classes are more liable for deletion then others:

(26)

LOW PROBABILITY <DELETION <HIGH PROBABILITY

CLASS	I	II	I/II PL	III SG	IV SG/PL
CM	*w	*r/y	*b	*b	*d
DPI	very high	high	high/middle	low	low
Control	very high	high	high	low	very low
IL[69]	very high	very high	high	low	low
Discrete	yes	yes	less	less/no	less/no

From this chart we may conclude that elements representing reflexes of the CM III **b* and CM IV **d* are in particular true candidates for petrified class markers, whereas any assumptions about PCM as reflexes of Class I and II should either be rejected or explained separately. For instance, Moor 1985:30 interprets Lezgi *wuĝun* 'to give' and *yaĝun* 'to throw' as consisting of the PCM **wu-* and **ya-*. The only sources for these "PCM" could be Class I and Class II. This, however, is incompatible with the above-mentioned conditions. Both verbs represent transitive concepts; accordingly, we should look for stereotypically deleted patients. The predication frame for both verbs shows that the prototypical semantics of the patient contain the feature [-hum]; cf.:

(27)

$\mathrm{THROW}_{V}\text{-}\ ((\Omega x_1\text{:}[+\mathrm{hum};+\mathrm{control}](x_1)_{AG}\ (\Omega x_2\text{:}[-\mathrm{hum};\ \pm\mathrm{anim}](x_2)_{PAT}).$

Thus, the petrification of class agreement should take place in the relation between the verb and its non-human patient, i.e., the patient should belong to Class III or IV. As we have no indication of a phonetic change **b* or **d* > *w* or *y* in Lezgi, we can assume that these elements do not represent PCM. Avar *réč:'ize* 'to throw,' however, fits well enough into this frame, as the initial *r-* can result from **d-*. Avar has lost Class IV completely, as we may see from forms like *réč:'ize,* a petrified reflex of this highly unmarked class. This procedure of checking whether the phonetic assumption of a PCM match with functional conditions is very helpful in those languages that have lost or changed the prototypical NC. Udi for instance has an interesting pair: *aq'sun* 'to hold' and *biq'sun* 'to take.' I have shown elsewhere (Schulze 1987.) that the *a/i*-variation reflects an old, proto-Lezgian type of aspect differentiation. Thus, we can reconstruct a common verbal stem **ĕq'-*, which in the case of *biq'sun* appears prefixed. This stem has its parallel in several other Lezgian languages, cf. Lezgi *q'un,* Tabasaran *biq'uz,* Rutul *x̂aq'n,* and Caxur *aq:'as* < proto-Lezgi **ĕq'(°)-* "to take.'[70] As initial **b-* is normally preserved in Udi, this element (segmented by internal reconstruction) can be equated to the CM III **b-,* which fits into the above-mentioned frame. In fact, the root structures *bVC-* and *dVC-* are quite prominent for verbs in Udi. Thus, we may assume that parallel processes of petrification concerning Class III and IV can be found in many Udi verbs.

Keeping in mind the above functional conditions, I want to return to the problem of possible PCM in nonverbal categories. The general assumption for nouns was that a noun containing a PCM must not have been a noun when the CM still functioned. But what was the category if we analyze such nouns? It seems probable that we have to deal with ancient adjectives or verbs in adjectival function. These adjectives (in attributive function) should have covered most of the prototypical semantics of their *determinatum*, so that the latter could have been deleted. Parallel processes are well documented in the history of languages. For instance, the Germanic word **beran-*, which became English *bear*, German *Bär* (as a *u*-stem, also Old Norse *bjǫrn,* Danish *bjørn* etc.) reflects the Indo-European adjective **bhero-* “brown’ (cf. Lithuanian *béras* ‘brown’), used for the concept <BEAR> by deleting its deteminatum ‘bear.’ The same can be true for several ECL lexemes. The easiest way to prove this would be to identify the root of a noun (without the assumed PCM) with a given verbal or adjectival stem (as was shown for Arči *mac’at:u(r)* ‘groom, bride’; cf. above). Unfortunately, external reconstruction in ECL is far from being at a point where it can offer us a stock of lexical items large enough to enable this identification. But in some cases it seems possible at least to propose a parallel interpretation, although the lexical background still is missing or less than obvious. Thus, the concepts for <DAY> and <SUN> in many languages are expressed on the basis of the same morpheme. Dargwa has only one word covering both concepts: *barħi*. But most of the ECL use two different lexemes which stand in a more or less clear relation, cf.:

(28)[71]

	sun	day		sun	day
Bežita	boq	wodo	Rutul	wireg̑	yěg̑
Cez	buq	g̑udi	Cax	urwirig̑	yěg̑
Xwarši	buq	g̑°ede	Kryz	wirag̑	yig̑
Hinux	buq	g̑°ede	Budux	wirag̑	yig̑
Xunza	boq	wede	Ag̑ul	rag̑	yag̑
Avar	baq’	q’o	Arči	barq	iq
Lak	barg̑	q’ini	Xinalug	enq’	q’ä
Lezgi	rag̑	yug̑	Udi	beg̑	g̑i
Tabasaran	rig̑	yig̑			

As can be shown by phonetic comparison, the basic stem common to all given lexemes was something like PEC **(ĕ)Gĕ.*[72] It probably was a verbal stem indicating 'be clear, bright,' as still preserved in the Arči stative verb *iq* 'to be bright.' There are several problems regarding the phonetics of the examples, especially the labialized forms in the Cezian languages (one may think of a secondary development); the starting point perhaps was **(ĕ)G°ĕ-*. The addition of the element *-de/-di* in the words for 'day' clearly is a Cezian innovation. The stem **(e)Gĕ-* obviously was used as an attributive adjective in stereotype collocations together with lexemes denoting the two objects which are characterized by this adjective and controlled its class marker, thus "sun" = *CM$_i$-(ĕ)Gĕ {NOUN$_i$[SUN]} and *CM$_i$-(ĕ)Gĕ {NOUN$_i$ [DAY]}. "Sun" would have been something like 'the shining one' and "day" something like 'the bright one.' The question as to how the two concepts are finally realized in the individual ECL is quite complex and beyond the problem to discussed here. The phonetic evidence in the Lezgian languages gives us **yĕG* for "day" and **w/bĕrĕG* for 'sun.' Thus, **yĕG* could only be analyzed as PCM-ADJECTIVE, if we assume that the CM II **y-* is older than **r-*, which is hard to explain in terms of proto-Lezgian phonetics. The other languages have suffixed forms for "day" and prefixed forms for "sun." For other reasons[73] we must assume that the prefixed forms are older than the suffixed ones. The labial stop in the northern languages especially reminds one of the CM III **b-*, which is quite liable to petrify according to the functional conditions given above. The reconstruction in the Cezian languages would lead to **bĕG°* for "sun," opposed to **G°ede* "day.' Thus **b(e)-* can easily be claimed to represent the CM III. The same will be true for Avar. The situation in the remaining languages, especially in the Lezgian, is much more complex. The reconstruction leads, as mentioned before, to a complex structure **w/bĕrĕG*. I would propose to treat the element **-rĕG-* as **bĕG°*, with the only difference being that we have to think of Class II as petrified. This root was augmented by another prefix, the nature of which remains unclear. The marginal languages Arči and Udi together with Lak may hint at the fact that the *b*-formant must be older than the **w-*, later lost in the Samur languages. Whether **b-* and **w-* had the same function (CM ?), whether **w-* is derived from **b-*[74] or whether they had nothing in common is still beyond our knowledge. We have to bear in mind that there probably were much more functionally important prefixes in proto-Lezgian than today (e.g., **rĕ-* could form *nomina instrumenti*, **mĕ- nomina actionis* from verbs), so that their identification as class markers perhaps may often be too precipitate. With this example, I want to show how difficult it is to determine the status of PCM with nouns. Only a consistent check through as much data as possible can reveal whether an element in question is a PCM or not. This check should always be conducted alongside with the above given functional conditions.

Up to now we have tried to describe the process of petrification in terms of the deletion of class-controlling nouns. This probably was not the only way. In particular, the fact that, for instance, in Udi (as well as in other ECL) transitive verbs with PCM still appear with patients hints at another procedure which was based not on deletion within stereotype collocations, but on the generalization of semantic properties. By this is meant that not only the high frequency of one noun in a proposition could be decisive, but also the high frequency of a specific, unmarked feature common to many nouns which were used with a verb. Let us take Udi *bot'esun* 'to beat off.' Its cognates in the other Lezgian languages are: Lezgi *at'un*, Tabasaran *hara-t'uz*, Rutul *ay-t'un*, Caxur *Ga-t'as*, Kryz *yi-t'äǯ*, Budux *yo(ro)-t'u*, Arči *at'as*, and Xinalug *t' ĭ̃*. The underlying proto-Lezgian stem **(ĕ)t' ̊-* probably meant "to slaughter." The root structure in Udi *bot'esun* < **bĕ-t' ẽ̆-* makes the interpretation of *b-* as a PCM highly probable. However, we cannot claim that the concept <TO BEAT OFF, TO SLAUGHTER> is specific enough that it could form a stereotype collocation. On the contrary, it refers to quite a number of semantically different objects in the world. In order to explain the way of petrification we have to ask what these possible objects have in common so that they frequently control Class III. Taking a look at the semantic features of class III (cf. (11)) leads us to the assumption that the concept of **(ĕ)t' ẽ̆-* frequently occured with [-hum;+anim] nouns in expressions with nouns denoting "socially relevant animals." It is just this inferential feature, which correlates perfectly with the concept given above, so that we can claim that **b-* petrified in Udi *bot'esun* because the verb very often was used with "animals that are or can be slaughtered." Interestingly enough, this process of generalizing objects of a certain class in their collocation with a specific verb led to a desemanticization of the verb (Udi "*to slaughter" > "to beat off"). Obviouly, the frequency in using *bot'esun* reduced its nuclear semantics and led to the necessity of introducing a new verb specific enough to replace the old concept <TO SLAUGHTER>. This was done with the help of Udi *šam-desun*. It seems clear that both the deletion of a noun and the generalization of common semantic features of nouns used with a specific verb can motivate the petrification of a PCM. The decision as to which process has taken place depends on the semantics of the verb in question. The more it denotes a state or an action affecting the protoptype nucleus of a noun the more likely deletion is. Or, vice versa, the more it denotes a state or an action affecting the vague, peripheral boundaries of a noun (in the above given example "all that is slaughterable"), the more likely generalization is. The process of generalization is of significant importance in those languages that still have NC. It can lead to a class change of nouns which originally did not belong to the class of the "frequent objects." Because of their usage with a specific concept which "normally" (i.e., frequently) was used with nouns of a typical class, they can enter this class as well.[75]

4. Conclusion. The above given can only be rudimentary. I did not want to discuss the numerous lexemes in ECL that have already been subject to the question of PCM.[76] I hope that many of them can automatically be eliminated from the list of PCM candidates by the method proposed in this paper. My proposal to introduce the concept of a Functional Diachrony into the discussion perhaps can help to systemize this discussion, but many things important in the domain of FD have had to be neglected.This is especially true for procedural aspects in FD, here regarding the question of how the functional system of noun classification has changed in the individual languages and proto-languages and whether these assumed changes have influenced the process of petrification. Introducing a "diachronic prototype" surely is disputable, but it can give us some information about the functional conditions of former morphosyntactic (sub)systems which were decisive for their development. Many things still remain to be done (and some can probably never be done). The most urgent task is to draw a complete picture of class-marking strategies in the individual ECL and to confront the lexical systems with respect to their class allocation, both in a semasiological and an onomasiological way. Finally, we have to discuss all the different procedural aspects of system change, especially the interaction of phonetic, functional, and semantic properties in order to get closer to the "gestalt" of the system of noun classification both in the proto-language(s) and in intermediate states.

REFERENCES

Andγulaʒe, N. 1968. *K'lasovani da p'irovani uγ vlilebis ist'oriis zogi sak'itxi iberiul-k'avk'asiur enebši*. Tbilisi (Meciereba).
Bokarev, E.A. 1959. *Cezskie (didojskie) jazyki Dagestana*. Moskva (Nauka).
Bokarev, E.A. 1961. *Vvedenie v sravnitel'no-istoričeskoe izučenija kavkazskix jazykov*. Maxačkala (DagUč).
Burčuladze, G.T. 1982. Reduplikacija i grammatičeskie klassy v lakskom jazyke. *VJa* 5:109-113.
Bußmann, H. 1983. *Lexikon der Sprachwissenschaft*. Stuttgart (Kröner).
Cercvaʒe, Il. 1954. Gramat'ik'uli k'lasebi andiuri k'ilok'aveb i. *IKJ* VI:427-434.
_____. 1965. *Andiuri ena*. Tbilisi (Mecniereba).
Charachidzé, G. 1981. *Grammaire de la langue Avar (langue du Caucase Nord-Est)*. Paris (Favard).
Deeters, G. 1963. Die kaukasischen Sprachen. *HdO* 7,1:1-79.
Dik, S.C. 1989. *The Theory of Functional Grammar I*. Dordrecht (Foris).
Džejranišvili, E.F. 1956. Okamenelye èlementy grammatičeskix klassov v glagol'nyx osnovax i otglagol'nyx imenax udinskogo jazyka. *IKJ* VIII:367-381.
Fähnrich 1984. Zur Stellung des Nachischen in der nachisch-daghestanischen Sprachfamilie. *Sprachen Kaukasiens*, ed. B. Wilhelmi. Jena:16-26.
Gadžiev, T.E. 1958. Sledy grammatičeskix klassov v lezginskom jazyke. *UZIIJaL* (Maxačkala) 5:217-226.
Geeraerts, D. 1983. Prototype Theory and Diachronic Semantics. *IF* 88:1-32.
_____. 1988. Cognitive Grammar and the History of Lexical Semantics. *Topics in Cognitive Linguistics* (ed. B. Rudzka-Ostyn). Amsterdam (Benjamins):647-677.

Helmbrecht, J. 1990. Funktionale Aspekte der Konnotation im Bereich tabubelasteter Rede. *Akten des 25. Linguistischen Kolloquiums, Paderborn* (in press).

Ibragimov, G.X. 1990. *Caxurskij jazyk.* Moskva (Nauka).

K'axaʒe, O. 1984. *Gramat'ik'uli k'lasebi dagest'nis monatesave enata šoris.* Tbilisi (Mecniereba).

Kerimov, K.R. 1984. K voprosu o pokazateljax grammatičeskogo klassa v strukture glagola xinalugskogo jazyka. *ElKJ* XI:149-157.

Kibrik, A.E. 1972. O formal'nom vydelenii soglasovatel'nyx klassov v arčinskom jazyke. *VJa* 1:124-131.

_____ and S.V. Kodzasov 1990. *Sopostavitel'noe izučenie dagestanskix jazykov. Imja, fonetika.* Moskva (IzdMosUniv).

Klimov, G.A. 1969. *Die kaukasischen Sprachen.* Übertragen von W: Boeder. Hamburg (Buske).

Lafon, R. 1952. Etudes basques et caucasiques. *Acta Salmaticensia. Filosofia y Letras 5(2).* Salamanca (Universidad).

Leskien, A. 1876. *Die Declination im Slawisch-Litauischen und Germanischen.* Leipzig (Hirzel).

Magomedbekova, Z.M. 1971. *Karatinskij jazyk.* Tbilisi (Mecniereba).

Magometov, A.A. 1965. *Tabasaranskij jazyk. Issledovanie i teksty.* Tbilisi (Mecniereba).

Moor, M. 1985. *Studien zum lesgischen Verb.* Wiesbaden (Harrassowitz).

Murkelinskij, G.B. 1949. Klassy i klassnye pokazateli v lakskom jazyke. *Trudy 2-j nauč. sessii Dag. nauč. bazy AN SSSR.* Maxačkala (DagUč):105-106.

Musaev, M. S.-M. 1968. *Leksika darginskogo jazyka. Sravnitel'no-istoričeskij analiz.* Maxačkala (DUP).

Osthoff, H. /K. Brugmann 1878. *Morphologische Untersuchungen auf dem Gebiet der indogermanischen Sprachen I.* Leipzig (Hirzel).

Rosch, E. 1975. Cognitive representation of semantic categories. *JeP* 104:192-233.

Šaumjan, R.M. 1936. Sledy grammatičeskix klassov (rodov) v agul'skom jazyke. *Jazyk i myšlenie* 7:219-226.

Schmidt, K.H. 1962. *Studien zur Rekonstruktion des Lautstandes der südkaukasischen Grundsprache.* Wiesbaden (Steiner).

Schulze, W. 1982. *Die Sprache der Uden in Nordazerajdžan. Studien zur Synchronie und Diachronie einer süd-ostkaukasischen Sprache.* Wiesbaden (Harrassowitz).

_____. 1987. Tracing Aspect Coding Techniques in the Lezgian Languages. *Papers of the 4th International Conference on the Non-Slvaic Languages of the USSR,* ed. H: Aronson and B. Darden. Columbus, Ohio (in press).

_____. 1988. Noun Classification and Ergative Construction in East Caucasian languages. *Studia Caucasologica I.*, ed. Fr. Thordarson. Oslo (NUP):251-274.

_____. 1988a. *Studien zur Rekonstruktion des Lautstandes der süd-ostkaukasischen (lezgischen) Grundsprache.* Wiesbaden (to appear).

_____. 1989. Zur Entwicklunsgdynamik morphosyntaktischer Subsysteme. *Actes du Vième colloque caucasien*, ed. C. Paris. Paris (to appear).

_____. 1990. Ergativität und Kongurenz. Versuch einer funktionalen Morphosyntax des einfachen Satzes im Lakischen. *REGC 4* (to appear).

_____. 1990a. The Udi language. *The Indegenous Languages of the Caucasus, vol 3.*, ed. R. Smeets and others. New York (Caravan) (in press).

_____. 1991. *Zur Entstehung der Personalflexion in einigen ostkaukasischen Sprachen* (manuscript, to appear).

_____. 1991a. Funktionale Aspekte des einfachen Satzes im Deutschen (in preparation).

_____. 1991b. "Control" in East Caucasian languages. Paper presented to the VII. International Conference on the Non-Slavic Languages of the Soviet Union. Chicago.

_____. 1991c. The "Nominativisation" of Udi. *REGC 5* (in preparation).

Talibov, B.B. 1960. O nekotoryx okamelenyx i poluokamelenyx èlementax v strukture lezginskogo jazyka. *Voprosy grammatiki, sbornik statej k 75-letiju akad. I. I. Meščaninova*. Moskva:168-177.

_____. 1980. *Sravnitel' naja fonetika lezginskix jazykov*. Moskva (Nauka).

Trubetzkoy, N. 1937. Zur Vorgeschichte der ostkaukasischen Sprachen. *Mélanges de linguistique et de philologie offerts a Jacques van Ginneken* (eds. J. Wils et al.). Paris (Klincksieck):171-178.

Watkins, C. 1966. Italo-Celic Revisited. *Ancient Indo-European Dialects*, ed. H. Birnbaum and J. Puhvel. Berkeley and Los Angeles (UCP):29-50.

Willée, G. and W. Schulze. 1983. Noch eine Variation über das Thema "Lemma." *SDv* 2:40-46.

Xajdakov, S.M. 1980. *Principy imennoj klassifikacii v dagestanskix jazykax*. Moskva (Nauka).

NOTES

[1] The earliest relicts of Dargwa are some glosses in an Arabic manuscript of a text by al-Ġ̣azzālī (1507). Lak glosses start at about 1600, the first longer text is a translation of al-Buṣīrī's Arabic Muhammad eulogy *al-Burda* (18th century).

[2] One exception may be Udi, which is often claimed to be the descendent of the language of Caucasian Albania, with which some 15 smaller inscriptions are associated (dated between the 5th and 9th centuries A.D.). Though there is a lot of historical, geographical, and even linguistic evidence supporting this hypothesis, we are, however, far from having deciphered these inscriptions coherently on the basis of Udi (for this, cf. Schulze 1982 and the references given there). The archives of the State Museum of Maxačkala contain, e.g., a stone cross (about 15 cm high), showing Old Georgian characters in a non-Georgian language (ca. 11th century). Perhaps

this inscription, together with some others, can be interpreted as a variant of Avar, corresponding to the fact that Christianity was well-installed in this area before the invasion of Islam. But no concrete research has been done on such sources up to now.

[3] He claims that "die materielle Gemeinsamkeit dieser Sprachen [i.e. the ECL, WSF] läßt sich verhältnismässig leicht erkennen: zur Zeit kann man annähernd vierhundert nominale und verbale Stämme sowie Affixe nachweisen, die allen Sprachen dieser Gruppe gemeinsam sind" (Klimov1969:53f.). It must be stressed, however, that this claim is much too optimistic. It refers exclusively to the works of Lafon 1952 and Bokarev 1961, both of which follow a hazardous method of "Kling-Klang" etymologies.

[4] In a linguistic area such as the Caucasus, which is characterized by a very low rate of areal divergency, it is often hard to decide whether a lexical correspondence is due to genetic facts or language contact. In order to resolve this problem, we should make reference to sound laws which, as it has been said, hardly have existed up to now. The situation seems somewhat better in smaller subgroups, e.g., in Rutulo-Caxur (ECL); cf. Schulze 1988a.

[5] The bracketing indicates the degree of genetic relationship.

[6] Together with the now extinct Xaputli.

[7] Cf. Deeters 1963:11. Fähnrich 1984 proposed a list of Lak-Čečen word correspondences. Trubetzkoy 1937 added Udi to this "substratum," which was to have influenced groups II/III and V, an assumption which, however, could not have been substantiated.

[8] Here an important role is played by the Lak genitive-ergative, which perhaps can be found as a relict in the originally prefixed ergative forms of the personal pronouns in the Nax languages (cf. Čečen *sō* 'I' (erg. *as*), *ħō* 'you' (erg. *aħ*) etc.). Prefixation of pronouns obviously was a technique proper to the genitive in PEC only.

[9] Hence, the Nax languages would play about the same role as Hittite in Indo-European.

[10] It has to be stressed that the top-down reconstruction should always coincide with facts emerging from assumptions about diachronic processes (bottom-up). Here, the second basic law of Indo-European reconstruction should be applied, as first formulated by Leskien 1876:VI: "Die Kriterien einer engeren Gemeinschaft können nur in positiven Übereinstimmungen der betreffenden Sprachen, die zugleich Abweichungen von den übrigen sind, gefunden werden." This is the most important indicator of genetic divergency, i.e., of the way internal dialectal processes in the proto-language lead to the formation of younger subgroups or languages.

[11] Together with the probably very important, but yet often neglected aspect of expressive sound change.

[12] Or "three dimensional," which would better correspond to the functional term "linguistic gestalt."

[13] A typical example would be the interpretation of the 1.SG. pronoun **eĝhom* as prefixed and ablauted form on the basis of a hypothetical **(e)me*, found in the verbal paradigm as the agreement marker for the 1.SG.

[14] As they were first noted by Rudolf Clausius (1822-1888) and statistically evaluated by Ludwig Boltzmann (1844-1906).

[15] Functional values may, however, get lost, if the world knowledge of a language community has changed, i.e., if it no longer needs these values in a communicative or social sense.

[16] This entropic tendency, which leads to a maximal balance between different functional categories, goes along with the "principle of least effort" (Zipf's law).

[17] In German terms "Paradigmenwechsel".

[18] It has to be stressed that I only deal with the basically monofunctional system of noun classification in ECL. There are, however, important relicts of a second, overt classification system, which was linked to the ergative morpheme (binary distribution of the feature [±control] leading to different semantic dichotomies). On this, cf. Schulze 1988.

[19] The examples contain only partial glosses. I have disregarded the infixation of the CM III (-w-) and the problem of Lak personal agreement (b-u-r). On this, cf. Schulze 1990.

[20] I have dealt with the syntactic properties of CM in Schulze 1988, the semantics have been discussed in Schulze 1989. Schulze 1988 briefly discusses the syntax of CM in the individual languages. It must be stressed, however, that a comprehensive, comparative study of CM syntax in ECL still is needed.

[21] Cf. Xajdakov 1980, K'axaʒe 1984 for a detailed discussion of the morphological inventories in the individual ECL. I will, in the following, disregard the tendencies, especially in some Lezgian languages (e.g. Caxur and Rutul), to develop new series of CM, partly based on phonetic processes involving CM (and leading, e.g., to ablaut phenomena in the verbal paradigm), partly representing true innovations (cf. the Arči CM *ħ-* (IV) with the two verbs of movement *ħ-erš:as* (IV) 'to run' and *ħ-erqās* 'to walk' (IV), which, however, do not occur in the perfect tense).

[22] Lower Andi has maintained the singular structure in the plural, but as it makes use of the same morphemes in the singular as in the plural, we might claim that Lower Andi does not have any plural codings by CM at all (on this, cf. Cercvaʒe 1954, Cercvaʒe 1965, Schulze 1988.).

[23] The elements "a," "b," "c," "d" are variables for the different types of morphological realization found in ECL. They do not represent concrete morphemes.

[24] Starting with Rosch 1975. The prototype theory can be regarded as the starting

point of cognitive linguistics as it is described, e.g., by Helmbrecht 1990 (with quite an extensive bibliography on prototype theories). On its application to diachronic semantics, cf., e.g., Geeraerts 1983 and 1988.

[25] The concept of diachronic prototypes has nothing to do with terms like *"etalon"* or "canonical," which propose a metalinguistic and abstract formulation of the construction in question and are used to evaluate language data according to whether they match these formulae or not. Naturally, there are mixed interpretations introducing empirical findings into and relativizing this idealistic approach.

[26] "Topic" in the sense of Dik 1989. In this paper I try to adapt the theory and formalism of Functional Grammar as developed by S. Dik, though it may be to restrictive to meet the conditions of a "vague," morphosyntactic diachrony as proposed here. That is why the formalism had to be modified in some points (faults in which, naturally, are my own responsibility). The tendency to use deictic elements as CM has been reintroduced by Xinalug, which uses its set of demonstratives to establish a new paradigm of CM (cf. *ē dä pši at:ĕ-dämä* (III) 'this horse is mine' and Kerimov 1984).

[27] "i," "j," etc., are referential indices. In this case, the pronoun would be coreferential with the noun.

[28] There may have been a parallel process in German; cf. the once marked position S1 (i.e., the position in front of the verb), which still is used in topic function with non-agents *(gestern*$_{top}$ *ging ich auf den Markt* '*yesterday* I went to the market.' Probably, this position once was marked with agents too *(*ich*$_{i/top}$ *ging* $_{(ichi)}$ *gestern auf den Markt)*, but became unmarked because of its high frequency in usage (cf. Schulze 1991a).

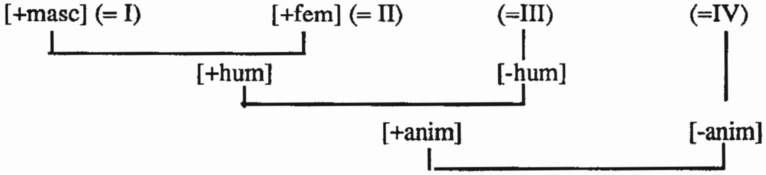
[29] I do not refer to the problem of the feature [±control] interacting with the ergative case here; cf. note 18.

[30] Each deviation from this number that can be found in the actual languages has to be explained on the basis of the four prototypical classes. The mechanisms to increase or reduce the number of classes follow certain general principles which are described in Schulze 1989.

[31] Though even this is not true for all languages. In Arči, for instance, nouns denoting certain human qualities may be encoded by means of Class IV, if the sex is unimportant, cf. *misgin qʕa‘* (IV) 'a poor one can' vs. *misgin qʕa‘* (I) 'a poor man came' and *misgin da-qa‘* (II) 'a poor woman came.' The same is true, e.g., for *lo* 'child,' which is class IV in the singular (when the sex is irrelevant to the speaker), but I/II in the plural, cf. Kibrik 1972:126.

[32] Cf. Schulze 1991b.

[33] A typical binary interpretation would be:



[34] The incorporation of non-human nouns in Class II represents an important factor in the tendencies toward splitting this group [±human] up , leading to a five to six class system in the Cezian languages Xunza and Bežita (Tladal dialect only).

[35] This cut-off point is of great importance for the development of the class semantics in Lak. A complex inferential process led to the incorporation of nearly all nouns [+fem; +hum] in Class III, whereas Class II is only used in highly marked constructions with a negative connotation; cf. Murkelinskij 1949:106, Schulze 1989. Still, *ninu* 'mother,' *š:ars:a* 'woman,' *q:ari* 'old woman,' and *susulu* 'widow' normally are encoded by Class II *(šars:a,* with a strong pejorative connotation according to my informants).

[36] Cf. Schulze 1988, Schulze 1990.

[37] This tendency to reduce the discrete properties of a noun in the plural matches perfectly well with restrictions in antipassives, for instance, in Bežita. Here, only [-discrete] nouns, i.e., collectives and plurals can be backgrounded with the help of the subjectivization of a then absolutive marked transitive agent, cf. *ist'iΛi-Ø ĝarλolca* 'the brother (ERG) boils water (ABS)' → *is-ØΛi-d ĝarλoldāc* 'the brother (ABS) boils (INTRANS) with water (INSTR).' An antipassive would not be possible in a proposition like *ist'i wahadi box̂oro-Ø soyra* 'the brother bought this horse (ABS).' A pseudo-antipassive like Lak *p:uØ- bawx̂unu ur ču-Ø* (father-ABS III-bought 3.SG-AUX horse(III)-ABS), however, can be translated as 'the father was a horse-buyer'; cf. Schulze 1990 for details.

[38] I will not treat the multiple proposals concerning CM as inflectional morphemes here. First, such a use of CM does not have anything to do with the process of petrification; on the contrary, it would lead to the formation of new paradigmatic structures. Second, many of these proposals are far from representing results of serious research; they often contradict simple evidence from the linguistic data (especially phonetics, noun and verb morphology, syntactic properties of the ergative construction, etc.). Thus, a complete revision both of the material discussed and its analysis becomes necessary. This shall be done in a separate study.

[39] There are only two classes in the singular (contrasting [±hum]) and one class in the plural. Probably Northern Tabasaran stopped the trend towards the elimination of NC because of contact interference from Lak and Dargwa, both languages with a quite stable system of NC and in close proximity to the language.

[40] Cf. Magometov 1965, Schulze 1991.

[41] Cf. Watkins 1966:31 for a detailed discussion of the term “negative innovation,” which contrasts with Leskien’s formulation of “positive Übereinstimmungen”; cf. note 10.

[42] Cf. Schulze 1982 and Schulze 199 1.c.

[43] It is hard to tell whether **r* or **y* was the morpheme for class II. **y* has to be reconstructed for proto-Avaro-Andi, **r* for proto-Lezgi. As *y* often emerges from **r* in initial position (and not vice-versa), we should assume that **r* was the underlying PEC realization and that the shift of **r-* > **y-* took place in times of dialectal differentiation of PEC. This assumption, however, is not yet supported by sound laws found between proto-Avaro-Andi and proto-Lezgi.

[44] This is not the place to enumerate all the studies which deal with PCM in this sense. Typical approaches are e.g. Anḍulaʒe 1968, Gadžiev 1958, Džejranišvili 1956, Talibov 1960, Šaumjan 1936. Xajdakov 1980 has the perhaps most comprehensive bibliography on CM research in ECL.

[45] It would be helpful to create a negative inventory of all those phonetic elements in the individual ECL which could, then, have never been a PCM. It could serve as a check list for corresponding assumptions.

[46] Cf. Burčuladze 1982. The Lak root *-ix̄:(i)-* has its regular correspondence in Dargwa *d-uq-is* ‘to separate’ *(d-* is a preverb). The question is whether *zix̂:in* can result from **d-ix̄:in* (IV) and whether the reduplicated form belongs in the same context (Burčuladze argues that in these cases reduplication serves to compensate for a lost CM in order to maintain a prefixed structure).

[47] I can disregard the variants in the dialects, as they do not shed essential light on the problem discussed here, cf. Schulze 1982:269, Schulze 1988a s.v. for a more extended discussion of proto-Lezgi **q’ °e-*.

[48] The underlying homorganicity of a labialized uvular/velar stop is well documented typologically; cf., e.g., the situation in p-Celtic, Umbrian, and Greek (I-E.. **kʷetu̯er-*, etc., ‘four‘ > Homeric *πίσυρες* Osc. *pettiur*, Welsh. *pedwar)*.

[49] This does not mean that numerals did not have prefixation in PEC. The proto-Lezgi forms reconstructed for “two” vs. “four” and “three” vs. “six” *(*q’ ° : *yĕq’ °* and **λe(b) : *rĕλe)*, for example, give us a prefix **r/yĕ-* probably indicating a dual morphology (“four” = “2-fold two,” “six” = “2-fold three”). But it does not have anything to do with CM!

[50] Cf. the definitions given in linguistic dictionaries. Here, I refer to Bußman 1983:297

[51] Cf. Schulze 1988a s.v. for the Lezgi data. An onomasiological list for the concept “daughter” in ECL is given by Kibrik and Kodzasov 1990:56. The PEC form is either **rĕčĕ* or (less probably) **yĕčĕ*.

[52] My transcription of laterals: Λ = voiceless spirant, λ = voiceless affricate.

[53] It is questionable whether Udi still had NC at this time or only emphasized the already feminine semantics of **či* by means of *x̂uni*.

[54] The Central Lezgian languages show analytic forms based on verbal nouns plus auxiliary, Rutul and Caxur use a common prefix (probably also found in the Kryz and Budux examples). The infinitive-marking techniques are described in the corresponding reference works.

[55] Cf. Schulze 1988a:202f. for the reflexes of *λ:' in the individual languages.

[56] Cf. Ibragimov 1990:116 with other examples. I only give the basic forms for the present tense and disregard suffixed morphemes based on the auxiliary and indicating personal agreement (I/II Class 1/2.Sg *-wor,* 3.Sg. *-or,* III/IV Class 1/2.Sg. *-wob,* 3.Sg *-ob,* etc.).

[57] Avar has retained three noun classes, Class IV being lost paradigmatically, though there are some important relicts of this class, such as *réč:'ize* 'to throw,' *rét:'ize* 'to dress,' *rék:'ize* 'to mount' (cf. Charachidzé 1981:30ff.). The procedures for detecting these PCM marked lexemes are parallel to those mentioned in the text concerning Udi *buq'sun.*

[58] This is not the place to discuss the terms "citation form" or "lemma" *in extenso* (reference can for instance be made to Willée and Schulze 1983). Ignoring the cognitive and functional aspects of these terms, I use them here to indicate the way lexical items are articulated by native speakers when being asked for correspondences in their language under "context free conditions."

[59] This is of significant importance for the problem of determining the functional value of a CM. Obviously, Bokarev's informants tended to associate nouns of Class II, i.e., mostly nouns denotationally marked [-mask; +hum], with the concept of <TO LOVE>. From there we can deduce that the second conceptual component of this verb, <TO WANT>, which is quite frequent elsewhere in ECL, does not play an important role in these three languages.

[60] The bracketing { } indicates petrification.

[61] *She-goat* can be regarded as a secondary form, where *she* is analogously added to the already feminine *goat* because of the masculine marked *he-goat.*

[62] The bracketing is used to indicate deleted lexemes.

[63] On the allocation of [+hum] nouns ("child" etc.) to this class, cf. 3.1.

[64] A source could be seen in Skrt. *snuṣā* 'daughter-in-law,' which has been simplified in forms such as *nus-* in the North, but to *sus-*, etc. in the Southern ECL. Arči *nus-* itself is a loan from a northern language, cf. Schulze 1988a s.v. Unfortunately, we do not have any areal or historical evidence for a Sanskrit contact with ECL; one might think of a loan from Mitanni, if we follow Gamkrelidze's view of a Mitanni migration along the Caspian coast line.

[65] Cervcaʒe 1965:328ff.

[66] Cf. Magomedbekova 1971:58. In rare instances, Karata analogously uses this technique with nouns of Class II, too, though we still have a genitive morpheme; cf. *yaši-l-ob qoča* (girl-GEN-III book(III)) 'the girl's book.'

[67] There seem to be rare exceptions; cf. the use of the German exclamation *bewahre!* 'God forbid!,' which shows the deletion of the stereotypical agent *Gott* (< *Gott bewahre!*).

[68] It seems logical that a *figura etymoligica* is the first candidate for deletion; cf. German *Wäsche waschen* > *waschen* 'to launder.'

[69] IL = degree of loss of information.

[70] The addition of Budux *surqu* 'to take' to this list by Talibov 1980:309 does not respect the formal diachrony (Budux *q* < **q'* is an ad hoc assumption). *Surqu* has its best parallels in Rutul *x̂ar-qan* and *fa-qĕn* 'to take.'

[71] I do not mention the Andian languages here, as they have developed a new set of lexemes for "sun" and "day".

[72] On this cf. Schulze 1988a s.v.

[73] At least proto-Lezgian, but probably also the other proto-languages were much more centripetal with regard to the order of agglutination then the present-day ECL. Suffix chaining, for instance, in Lak and Avar surely are younger phenomena. The centripetal "look" of PEC relates it in an interesting way to the West Caucasian languages, which is perhaps not merely a typological accident.

[74] In fact, a development *b-* > *w-* is possible in many ECL. However, there are hardly any word correspondences in the Lezgian languages which could prove this shift; cf. Schulze 1988a:60f.

[75] This is probably one of the reasons why there is a strong tendency in NC systems towards homorganicity. This entropic tendency leads to an equilibrated system of semantic allocations, the result of which may be, for instance, a purely diptotic system as proposed in note (33) or to a merely grammatical gender (cf. the tendencies in Avar, the NC system of which often is regarded as a gender system).

[76] I also disregard the very important question of PCM with personal and demonstrative pronouns. This will be subject of an independent study.

IBN RUSTA AND BAKRĪ ON THE OSSETIANS AND THE APKHAZ

David Testen
University of Chicago

Ibn Rusta in *Kitāb al-ʹaʻlāq al-nafisa* mentions two peoples, the *Ṭwlās* and the *Lwġr,* inhabiting the mountain region bordering on the land of the Khazars. In discussing the location of these peoples he observes, "The Khazar country is a vast land, one of whose sides adjoins a huge mountain, and this is the mountain at the farthest end of which live the *Ṭwlās* and the *Lwġr,* and this mountain stretches to the land of Tiflis."[1] Similar statements, with varying forms of the second name, may be found in the works of other, later writers: ʻAwfī ("The land of the Khazar is a long and broad realm, and on one side is a great mountain, and from the limit of the mountain two types of Turk come down, one people of which are called *Ṭwlās* and the other one *Kwġr,* and this mountain adjoins the land of Tiflis"), Šukrullāh b. Šihāb ("And in the middle of the Pecheneg and the Khazar is a long and broad realm, and on one side is a great mountain, and at the [limit] of that realm two types of Turk come down, one is called *Ṭwlas* and one *Krġrh*"),[2] and the anonymous *Ḥudūd al-ʻālam* ("ṬŪLĀS, LŪGH.R (?), two regions of *(az)* the Khazar (country). The people are warlike and have great numbers of arms") (Minorsky 162).

In the first of these two names Marquart has identified the element *-ās* as *Ās(i),* which appears to have been the self-designation of the Alans, or perhaps of a prominent subgroup of the Alans. This name is attested in the *Jas-i* of the Old Russian chronicles, and it replaces *al-Lān* (and its various alternative spellings) in later (Mongol-period) Muslim sources to designate the Alans. In Ossetian, the modern descendent of the language of the Alans, the term *As(s)y* refers to the region of the Balkar in the north Caucasus, inhabited by the Ossetians before the advent of the present Turkic-speaking population. Other toponyms based upon this name, coming from areas formerly occupied by the Alans, include *Jászberény* in central Hungary, *Iaşi* in eastern Romania, and, perhaps, *Astraxan* at the mouth of the Volga. By Marquart's reading, therefore, the *Ṭwl-ās* were a specific group among the Alans; the structure of their name calls to mind the tribe of the *Dxsās,* said by Ibn Rusta to be the noblest tribe of the Alans. This name has been emended by Minorsky to **Rxsās* and read **Ruxs-ās* ("Light-Ās"—cf. Ossetian *ruxs* 'light' < Iran. **rauxšna-*), a form conspicuously reminiscent of the name of the *Roxolanoi* of the Greek and Latin authors (Minorsky 445 fn. 5)

In the first element of the name *Ṭwlās* Marquart attempted to find *Dula,* the name given to an Alan prince in an early Magyar chronicle (Marquart 172). A more plausible interpretation, however, is that of Minorsky, who sees in *Ṭwl-* the name of the *Tual-tä,* the Ossetian tribe inhabiting the Ardon River valley, whose name has often been misused as a general cover term for the Ossetians living south of the Caucasus range (Minorsky 457).

It is Minorsky's view that this interpretation of the *Ṭwlās* may be reconciled with Ibn Rusta's geography by bearing in mind that the term "mountain" in Arabic *(jabal)* or Persian *(kūh)* may refer to either a single peak or to a mountain range. He therefore suggests that the *jabal* in question, which marks the edge of the Khazar domain, must be the Caucasus range. Once the Arabs took control of Albania and occupied Derbent (Bāb al-'Abwāb), the mountains came to form the southern limit of Khazar power. Since the Khazars were based at the eastern end of Caucasia, the mountain's "farthest end" *('aqṣā-hu)* to which Ibn Rusta refers should be the western end of the range, on the Black Sea.

The name of the people paired with the *Ṭwlās,* as has been seen above, shows a good deal of variation among the sources: لوغر *Lwġr* (Ibn Rusta, *Ḥudūd al-'ālam*), كوغر *Kwġr* ('Awfī), كرغر *Krġr* (Šukrullāh). Marquart reads these as distortions of *اوغز **'wġz,* *Awġaz, which he takes to be a variant of *'bxāz,* Apkhaz, the well-known people settled on the Black Sea coast, whose name was frequently extended in medieval Muslim sources to refer to Western Georgia in general. In this case, however, Minorsky believes that the reference is to the Apkhaz *sensu stricto,* basing this belief on the account of Bakrī which we will address shortly. This interpretation of Ibn Rusta's account, therefore, appears to locate the *Ṭwlās* and the *Lwġr* at the western end of the Caucasus, to the west of the Georgians of the Tbilisi region.

There are problems with this reading, however. Most important is that, although it is clear that the ancestors of the Ossetians formerly lived much further to the west than their descendants do, the modern Tual group is not found in the west at all, but in the central highlands of the Caucasus, to the north of the central pass. Abaev, challenging the use of the name *Tual* by early Ossetianists to refer to the southern Ossetians, describes their territory in the following terms:

> ... Vs. Miller writes in the "Osetinskie ètjudy": "Tual is spoken by the southern communities, the Tuals, i.e. the Ossetians on the far side of the Caucasus, adjoining the Georgians." If one is to call "Tuals" those for whom the Ossetians themselves use the name, then Vs. Miller's claim is fundamentally mistaken: the Tuals live not on the far side, but on this side, of the "Caucasus," i.e. the mountain range, and, of course, nowhere do they border upon the Georgians. The name "Tuals" is used primarily of the inhabitants of the upper mountain regions of Northern Ossetia in the highlands of the River Ardon, beginning from the Kassar ravine up to the mountain range... (Abaev 495)

Whether or not the Tual live south of the Caucasus, it is out of the question that they should be found in the far west, and thus there is no obvious reason that they should have been associated so closely with the Apkhaz by Ibn Rusta and the other writers.

No less problematical is Ibn Rusta's insistence on connecting these two peoples with the edge of the Khazar domain. Although the western extent of Khazar power in northern Caucasia is not documented, it clearly never reached as far west as the regions associated with the Abkhaz. While it may be argued that it is possible to read Ibn Rusta in such a way that it is the "mountain" which is contiguous to the Khazars and not the peoples themselves, there is no question that the *Ḥudūd al-ʿālam* lists the *Ṭwlās* and the *Lwġr* as provinces subject to the Khazars. Minorsky admits to some puzzlement at this problem in his commentary to the *Ḥudūd*.

> ... [Ibn Rusta] only says that at the farthest end of the mountain near which lay the Khazar land, lived the *Ṭulās* and *Lugh.r,* whereas our author [i.e., the author of the *Ḥudūd*—DT] makes of

> the latter "two districts of the Khazar". It is true that in the seventh century the Khazars penetrated down to Tiflis through the central Caucasian pass but the western Caucasus was hardly ever under Khazar sway. Our author's mistake may be somehow connected with the frequent confusion of خزر *Khazar* with جرز *Jurz* "Georgians". I.R.'s detail on the mountain "stretching to the land *(bilād)* of Tiflis" is perhaps a hint of some mention of the Jurz in the original source. (Minorsky 457)

The clearest evidence which has been adduced in support of the equation of the *Lwġr* and the Apkhaz is to be found in Bakrī's *Kitāb al-masālīk wa-l-mamālīk.* His account contains a passage which is a close analogue to Ibn Rusta's description of the Khazar frontier: at the end of the Khazar realm is a great mountain, at the far end of which lies the land of Tiflīs, beyond which is Armenia. Ibn Rusta's *Ṭwlās* and *Lwġr* are absent at this point in Bakrī's text. In his chapter on the Magyar, however, Bakrī adds a note on two mountain peoples who have been taken to be the same as the *Ṭwlās* and the *Lwġr* of the other writers.

> ... One border of the Magyar land reaches the land of the Romans, and at the end of their region in the vicinity of the steppe is a mountain inhabited by a people called اين, who possess horses and cattle and fields... Beneath this mountain on the shore of the sea is a people called ارغونه who are Christians and border upon the lands of Islam belonging to the lands of Tiflīs. It is the beginning of the region of Armenia. This mountain stretches to the land of Bāb al-'Abwāb and reaches the land of the Khazars.... (Marquart 31, 173).

Marquart claims that the اين and the ارغونه of Bakrī are to be equated with the *Ṭwlās* and the *Lwġr,* respectively. He therefore maintains that the former name is to be read as a corruption of the name اس *Ās,* which we have seen above in the compound *Tual-ās*. This emendation entails taking the *yā'* to be originally the tooth of a *sīn,* and the final *nūn* to be its tail; the pointing, one must assume, is all secondary. He therefore believes that the *'yn* are not the Ossetians but the Alans, well known as a major power in the central Caucasus.

He deals with اوغونه _'wġwnh_ similarly, emending this form to *اوغزيه _*'wġzyh,_ i.e., *Awġaz-iyya _(Apkhaz-iyya),_ with the suffix _-iyya,_ a familiar gentilic formant. Bakrī thus provides the important statement that the *Awġaz-iyya (=*Awġāz) lived on the coast. The Alans (i.e., the _'yn_), if they are to be closely connected with the *Awġaz(iyya), as Marquart's reading of Bakrī would have it, must once again be located far to the west of the present territory of the Tuals (Marquart 172-6).

Minorsky does not disagree with Marquart's interpretation of the _'wġwnh_ as the Apkhaz, but he does challenge the reading of _'yn_ as _*Ās_.

> Although the forms _*D.khs-Ās_ and _Ṭwl.-Ās_ occur already in Ibn Rusta as the names of special tribes, the pure form _Ās_ as referring to the Alāns in general appears only in Mongol times. Moreover Bakrī's description of the اين lacks the characteristic features of the Alān. Even the combination of اين with the _*Aughaziya_ suggests that Bakrī has in view the particular clan corresponding to Ṭwlās... (Minorsky 458)

We can find difficulties with both this interpretation and Marquart's, since, as we have noted above, they both compel us to locate the _Ṭwlas/'yn_ and the _Lwġr/'wġwnh_ in the far west, rather than in central Caucasia as our knowledge of the present position of the Tual Ossetians would dictate. Under the circumstances, I would like to suggest going back to re-read Ibn Rusta, keeping in mind both Minorsky's very attractive equation of the _Ṭwlās_ with the Tual Ossetians and their present location on the Ardon River north of the center of the Caucasian range.

We may note first of all that this location places them quite close to where we must assume lay the western limit of Khazar domination. We know that the sphere of the Khazars' influence ended somewhere in the central Caucasus, even if the precise location is not defined in the material available. The _Ḥudūd al-'ālam,_ we have seen, states that the Tual were a province of the Khazar, and, since we have no reason to doubt it in this point, we can assume that the Tual were, at least at some point in their history, subordinated or allied to Khazar power, perhaps as vassals in charge of the Khazar frontier.

The crux of the geography of Ibn Rusta's account lies in correctly identifying the "great mountain" at the end of the Khazar realm. Both Marquart and Minorsky assumed that this must be a reference to the Caucasus range, which formed the southern limit of the Khazars. This reading, we have seen, forces us to place the Tual, incorrectly, far to the west of where they should be located, since the "furthest end" of the range could only be the end furthest from the Khazars.

Must we assume that the "great mountain" to which Ibn Rusta refers must be the Caucasus, however? If we begin by trusting the equation of the *Ṭwlās* and the Tual, and assuming that the Tual lived at the time at roughly the same place that we find them today, we find that we have another option in reading this passage: *jabal* here does not mean "mountain range" but "mountain," and the "great mountain" in question is not the Caucasus as a whole but the highest peak in the central Caucasus, Mount Qazbeg . With a height of 5,047 meters, Qazbeg is clearly a *jabal 'að̣īm,* and its location relative to the Tual Ossetians accords well with Ibn Rusta's description, since the Ardon River valley, where the Tuals are centered, lies directly to the west of the mountain, at its "furthest end" from the point of view of one travelling from the Khazar lands in the east.

Interpreting Ibn Rusta's mountain as Qazbeg enables us to revise our understanding of the nature of this passage. Rather than finding in the text a hodge-podge of Ossetians, Apkhaz, and Georgians, thrown together with no apparent comprehension of the proper geography involved, I think that we have in Ibn Rusta's account a quite faithful description of what a traveller would have encountered if he were to journey through the eastern Caucasus at that time. Setting off from the Khazar heartland on the Caspian coast, the traveller seeking the lands of Islam would have two possible routes across the Caucasus. Either he could head south along the coast to Bāb al-'Abwāb and, beyond it, Albania and Azerbayjan, or he could head west through northern Caucasia. If he followed the latter route, he would reach the limit of Khazar power, we would suggest, in the vicinity of Qazbeg and the pass leading southward. Beyond Qazbeg—i.e., to the west—lay the lands of the Tual Ossetians. Turning south, however, brings him to the Darial Pass (Bāb al-lān, the modern Krestovyj pereval), beyond which he found Transcaucasia and the road heading southward to Tbilisi, probably tracing a route similar to that of the modern Georgian Military Highway.

I suggest, therefore, that in light of the fact that Ibn Rusta pinpoints the Tuals quite clearly in the central Caucasus we should not interpret the *Lwġr/Kwġr/Krġr* as the Apkhaz of the far west, since the geography of this reading of the text brings us nowhere near that area. The equation of this people with the Apkhaz, as we have seen, poses major problems if we assume that the *Ṭwlās* occupied the same territory that they do today: why do Ibn Rusta and the other writers so consistently make a link between the central Caucasian Tual and a people separated from them by half the length of the Caucasus? Why, moreover, would the *Ḥudūd al-ʻālam* claim that the Apkhaz are a province of the Khazars? If we allow ourselves some slight freedom to interpret the various forms of the name provided by the sources, we can easily arrive at a name which we could expect to find associated with the Tual, namely that of the *Digor* or Western Ossetians.

The paleographical interpretation of *Lwġr* (etc.) as **Dġ(w)r* poses no major problems. In the initial *kaf* of ʻAwfī's كوغر *Kwġr* and Šukrullāh's كرغر *Krġr* we may have a remnant of an earlier *dāl*, and the initial *lām* of Ibn Rusta's لوغر *Lwġr* may represent a trace of an earlier Arabic definite article. Perhaps the name in the original source had the form *دغور **Dġwr*, or perhaps *الدغور **al-Dġwr;* the emendations which we would have to make in either case are certainly no more major than those which Marquart and Minorsky assume in order to find the hypothetical **'wġz*. We are also, incidentally, no longer obliged to posit an alternative form for a name well attested under the familiar shape *'bxāz*.

We find the major objection to this reading of *Lwġr* (etc.) in Bakrī's description of the *'wġwnh,* where it is made clear that the latter people live on the coast. We should point out, however, that we have no real reason whatsoever to assume *a priori* that the two names *Lwġr* and *'wġwnh* actually refer to the same people. Even if we accept the assumption of Marquart and Minorsky that the *Ṭwlās* and the *'yn* are identical, we are hardly compelled to make a similar assumption about the names accompanying each of these. When all is said and done, we know only that the people called *'wġwnh* by Bakrī were Christians, lived on the coast, and were near the Georgians. Perhaps the name is indeed to be equated with that of the Apkhaz, whether the Abkhaz proper or the western Georgians who were given that name by the medieval geographers.

On the other hand, the possibility that the term *'wġwnh* referred to the Digor Ossetians cannot be ruled out. The present geographical location of the Digor does not render this possibility impossible, since it is clear that the Ossetians formerly extended much further west than the limits of their present distribution. The oldest known inscription in Ossetian, dating from the eleventh or twelfth century, was found on the Zelenčuk River, in the Karačai-Čerkesskaya Autonomous Oblast' of the RFSFR, and Abaev has produced a good number of toponyms showing that there was a significant Ossetian presence throughout much of the Kuban valley (Abaev 271-90), and the *Ḥudūd al-'ālam* describes the territory of the Circassians as "a land of the Alan on the coast of the Gurz [i.e., Black] Sea" (Minorsky 161). These facts imply that the ancestors of the Ossetians held sway over territory a great deal to the west of the land which they presently occupy.

Whatever the proper reading of *'wġwnh* may be, I would like to suggest that the identity of the people accompanying them in the account of Bakri is much less problematical. We have seen that Marquart wished to see in the اين *'yn* an alternate name for the Alans, and recommended emending the form in the text to **'s*. Minorsky objected to this interpretation by indicating that this people must be the same as the one to which Ibn Rusta gave the name *Ṭwlās,* the Tual tribe of the Ossetians, but he did not address the question of the name *'yn* which Bakrī provides. A simple way of reconciling Ibn Rusta's *Ṭwlās* with Bakrī's *'yn* is to make the minor emendation of replacing *nūn* with *rā'*. The resulting اير **'yr* would be an obvious way of rendering the Ossetian self-designation *Ir,* used today by the Tual as well as by the other East Ossetians. The appearance of **'yr* in the place of *Ṭwlās* would thus be a case of identifying the same group at different levels of reference: *Ṭwlās* represents the more immediate, tribal designation of the people in question, while **'yr* represents their broader, ethnic affiliations. This differentiation of levels may have been conditioned by oppositions existing within each text, if our conjectures about the different natures of the *Lwġr* vs. the *'wġwnh* are justified. It would make sense to find a tribal designation used to distinguish this people from the Digor but an ethnic term used when they are juxtaposed to the non-Ossetian Apkhaz.

The accounts by Ibn Rusta and Bakrī which we have been examining each provide us with an intriguing connection between the Ossetians of their day and the Ossetians of ours. If our interpretation of Bakrī is correct, we find that a certain group of Ossetians in the middle ages, just like their descendents in the twentieth century, are identified simultaneously as *Tual* and as *Ir*. In the case of Ibn Rusta's text, our revised interpretation of the geography involved has made it clear that the Tual Ossetians of the tenth century lived in virtually the exact spot that today's Tual inhabit, the Ardon River valley just to the west of Mount Qazbeg. The continuity across time which these texts indicate in the case of a small people like the Ossetians, confined, so to speak, to the marginalia of history, thus stands in striking contrast to the ephemerality of the major powers like the Arabs, the Khazars, and the Alans, on whose account the maps are drawn.

REFERENCES

Abaev, V. I. 1949. *Osetinskij jazyk i fol'klor*. Moscow-Leningrad: Akademija nauk SSSR.

Ibn Rosteh (ed. M. J. de Goeje). 1967. *Kitâb al-A'lâk an-Nafīsa* VII (Bibliotheca Geographorum Arabicorum VII). London: E. J. Brill.

Marquart, J. 1903. *Osteuropäische und ostasiatische Streifzüge*. Leipzig: Dieterich'sche Verlagsbuchhandlung.

Minorsky, V. (trans. and comment.). 1937. *Ḥudūd al-'Ālam 'The Regions of the World' : A Persian Geography 372 A.H.—982 A.D.* (E. J. W. Gibb Memorial Series (New Series) XI). Oxford: Oxford University Press.

NOTES

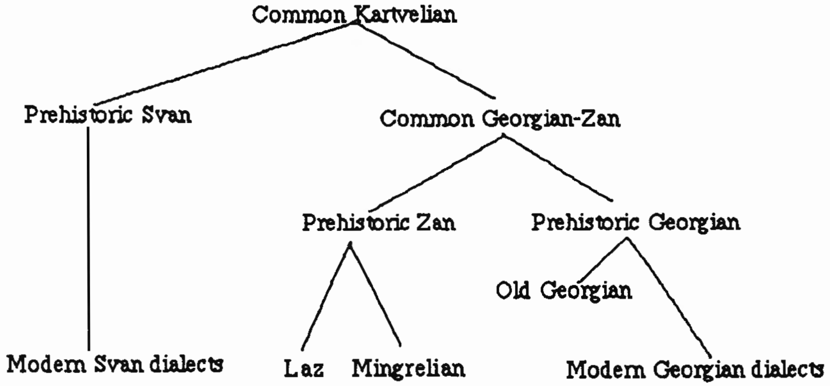
[1] “... wa-bilādu l-xazari bilādun ‘arīḍatun yattaṣilu bi’iḥdā janabātihā jabalun ‘að̣īmun wa-huwwa llaðī yanzilu fī ’aqṣāhu ṭwlās wa-lwġr wa-yamuddu hāðā l-jabalu ‘ilā bilādi tiflīsa...” (Ibn Rusta 139).

[2] Šukrullāh: “...wa-miyāna-yi bajanak wa-xazar wilāyatīst ṭawīl wa-‘arīḍ wa-bar yak ṭarfī kūhī ast buzurg wa-dar bayānān-i ān wilāyat dū naw‘ turk fīrod āyand yakī rā ṭwlās wa-yakī rā krġrh xvānand...”; ‘Awfī: bilād-i xazar wilāyatī-yi ṭawīl wa-‘arīḍ ast wa-bar yak ṭarf o kūhīst buzurg wa-az pāyān-i kūh dū naw‘ turk faro āyand ki yak qawm rā ṭwlās xvānand wa-dīgarī rā kwġr wa-īn kūh ba-zamīn-i tiflīs paywasta ast...” Unfortunately, neither of these texts is available to me. I have presented here the passages as they are found in Marquart (173, fn. 3).

THE CATEGORY OF NUMBER IN COMMON KARTVELIAN

Kevin Tuite
Université de Montréal

1. Kartvelian. The Kartvelian (South Caucasian) family comprises three languages: Georgian, Zan (Laz-Mingrelian) and Svan. Svan is the outlying member of the family; it is believed to have diverged from the rest of the family in the 2nd millenium BC. The nomenclature used here comes primarily from Harris [1985:6]. The protolanguage will be referred to as Common Kartvelian (CK), and the ancestor of Georgian, Laz and Mingrelian will be called Common Georgian-Zan (CGZ). I will refer to early, unattested stages of the individual languages as Prehistoric Georgian, Prehistoric Svan, and Prehistoric Zan.



In this paper I will examine the systems for coding number features in the Kartvelian languages and present a reconstruction of the system of number in Common Kartvelian. I hope to demonstrate in particular that Common Kartvelian had a well-developed system of verbal-plurality marking, which was indicated by morphemes common to both verbal and nominal morphology.

1. 1. Person and number marking in the Kartvelian languages. All three members of the family employ two sets of person/number agreement markers, shown in Table 1:

TABLE 1
Kartvelian person/number agreement markers

Set S ("subject") affixes

	Old Geo.		*Mod. Geo.*		*Zan*			*Svan*	
1sg	v-	-Ø	v-	-Ø	v/b-	-Ø,r,k		xw-	-Ø
1pl	v-	-t	v-	-t	v/b-	-t,rt	*excl:*	xw-	-(š)d
							incl:	l-	-(š)d
2sg	x-	-Ø	Ø-	-Ø	Ø-	-Ø,r,k		x-	-Ø
2pl	x-	-t	Ø-	-t	Ø-	-t,rt		x-	-(š)d
3sg	-s,a/o,n		-s,a/o		-s,u,n			(l)-	-s,<a>,Ø
3pl	-an,en,es,ed		-an,en,es,nen		-an,es,nan			(l)-	-x

Set O ("object") affixes

	Old Geo.	*Mod. Geo.*		*Zan*		*Svan*	
1sg	m-	m-		m-		m-	
1excl	m-	gv-		m-	-t,an,es,nen	n-	
1incl	gw-	gv-		m-	-t,an,es,nen	gw-	
2	g-	g-		k/g/r-		ǰ-	
2pl	g-	g-	-t	k/g/r-	-t,an,es,nen	ǰ-	-x
3	x-	Ø,s/h/Ø-		Ø-		x-	
3pl	x-	Ø,s/h/Ø-	-t	Ø-	-t,an,es,nen	x-	-x

In the following sections, I will present a reconstruction of the Common Kartvelian system of person/number agreement, largely based upon the work of Klimov [1964], Oniani [1978] and Schmidt [1982].

1.1.1. Georgian-Zan number agreement. The plural suffix associated with S1 (Set S, 1st person) and S2 marking is *-t* in Georgian and Zan, and one can assume an ancestral **-t* in Common Georgian-Zan. In both Georgian and Zan, the S3 markers are portmanteau morphemes, indicating tense and mood as well as person and number. The relation between tense/mood and S3 marking in Old Georgian was the same for both major verb classes (Class A ("active") and P ("passive")[1]). Since that time there has been a significant realignment of this pattern in literary Georgian and most modern dialects [Oniani 1979; Tuite 1988 §2.3.3.2]. Four groups of tense/mood forms ("screeves") can be discerned, each with its own set of S3 markers. The forms reconstructed for Common Georgian-Zan are in most cases those given in Klimov's etymological dictionary [1964].

TABLE 2
Cognate S_3 (Set S, 3rd person) suffixes in Old Georgian and Zan

screeve group		*3sg*	*3pl*
A. Present/permansive	OGeo:	-s	-en/an
	Zan:	-s	-an
	CGZ:	*-s	*-en
B. Conjunctive	OGeo:	-s	-n
	Zan:	-s	-n
	CGZ:	*-s	*-n
C. Past indicative	OGeo:	-a	-es
	Zan:	-u	-es
	CGZ:	*-a	*-es
D. Iterative/present	OGeo:	-n/ed	-ed
	Zan:	-n	-nan
	CGZ:	*-n	*-ed(?)

The feature composition of the agreement affixes of early Old Georgian (5th-8th centuries) was quite different for Set S and Set O, though each set was internally symmetric (Table 3).

TABLE 3
Person agreement system in early Old Georgian
[pl = plural; sp = speaker, ad = addressee]

Set S				Set O			
v-	<1, -pl>	v- -t	<1, +pl>	gw-	<+sp, +ad>	m-	<+sp, -ad>
x-	<2, -pl>	x- -t	<2, +pl>	g-	<-sp, +ad>	x/Ø-	<-sp, -ad>
-s/a/n	<3, -pl>	-en/es	<3, +pl>				

Comparison with Svan makes it reasonable to assume a similar feature composition for Common Georgian-Zan. Note that at this stage there was no feature of number (<±plural>) in Set O. The presence of specifically plural Set O affixes in Georgian and Zan is due to separate innovations. The plural suffixes which were originally restricted to the Set S system were extended to the Set O system to code the plurality of discourse-prominent NPs (1st/2nd person NPs and subjects — for detailed documentation see Tuite [1988 §§3-8]). In Zan, the choice of suffix to mark plurality in Set O is dependent on tense and mood (as with S$_3$), and also on the person of the morphological subject. This reanalysis of the function of the plural-agreement suffixes, along with the loss of the old opposition between inclusive and exclusive 1st person (already underway by at least the 7th century [Met'reveli 1978; Šanije 1982: 74; Sarǰvelaje 1984: 407]) resulted in the development of a similar feature structure for the two sets of agreement markers.

1.1.2. Common Kartvelian number agreement. The plural suffixes associated with S$_1$ and S$_2$ marking are *-d* and *-šd* in Svan [Deeters 1930:57-60]. This latter suffix is restricted to the copular verb in the Upper Bal dialect [Topuria 1967:9]:

xw-i 'I am'	xw-i-šd 'We$_{\text{excl}}$ are'
	l-i-šd 'We$_{\text{incl}}$ are'
x-i You$_{\text{sg}}$ are'	x-i-šd 'You$_{\text{pl}}$ are'
l-i 'S/he, it is'	l-i-x 'They are'

Klimov [1964:67-8] proposes _*-d_ as the Common Kartvelian ancestor, but notes that a Common Kartvelian _*-(s$_1$)t_ could also account for the two Svan forms as well as Georgian-Zan *-t* (cf. Gamkrelidze and Machavariani [1982:49,80]). It is unclear whether Common Kartvelian had an inclusive/exclusive opposition in the Set S system, as there is in some Svan dialects (Deeters [1930:27-28]; Melikišvili [1977]; Oniani [1984]), and has been reconstructed for Prehistoric Svan. Oniani [1978: 229-230] considers the O$_{1\text{exclpl}}$ (Set O, exclusive 1st person plural) prefix *n-* to be an innovation in the Prehistoric Svan period, encroaching upon the semantic range of O$_{1\text{excl}}$ *m-* , which was inherited from Common Kartvelian (Table 4):

TABLE 4

1. Person agreement system in early Prehistoric Svan

Set S					Set O	
xw-	<+sp, -ad, -pl>	xw-	-(š)d	<+sp, -ad, +pl>	m-	<+sp, -ad>
		l-	-(š)d	<+sp, +ad, (+pl)>	gw-	<+sp, +ad>
x-	<-sp, +ad, -pl>	x-	-(š)d	<-sp, +ad, +pl>	ǰ-	<-sp, +ad>
(l)-	<-sp, -ad, -pl>	(l)-	-x	<-sp, -ad, +pl>	x-	<-sp, -ad>

2. Person agreement system in later Prehistoric Svan

Set S					Set O	
xw-	<+sp, -ad, -pl>	xw-	-(š)d	<+sp, -ad, +pl>	m-	<+sp, -ad, -pl>
					n-	<+sp, -ad, +pl>
		l-	-(š)d	<+sp, +ad, (+pl)>	gw-	<+sp, +ad>
x-	<-sp, +ad, -pl>	x-	-(š)d	<-sp, +ad, +pl>	ǰ-	<-sp, +ad>
(l)-	<-sp, -ad, -pl>	(l)-	-x	<-sp, -ad, +pl>	x-	<-sp, -ad>

Since the evidence for an inclusive/exclusive opposition in Set O for Common Kartvelian is very compelling, it is reasonable, on the grounds of symmetry, to propose that the features <±sp, ±ad> were also used to structure the Common Kartvelian Set S affix system. The presence of distinct inclusive and exclusive S1 forms would then be a conservative element of Svan morphology, and not an innovation.[2]

Several proposals have been advanced concerning the origins of S3 marking. The only S3 markers found in all Kartvelian languages are *-*s* (3sg conjunctive [Group B]) and *-*a* (3sg aorist/imperfect) [Klimov 1964: 41-2,161; Boeder 1979: 471 note 22; Harris 1985: 397,402].[3] The other markers can be traced back to the Common Georgian-Zan period at best. The likely explanation for this state of affairs is either that portmanteau S3 suffixes such as those of Georgian and Zan also existed in prehistoric Svan, but later were eliminated — a change comparable to that now underway in the southwest Georgian dialects [Tuite 1988 §6; J̌orbenaje 1989] — or that screeve-dependent S3 marking did not develop until the Common Georgian-Zan period. Oniani [1978], Klimov [1979], and Schmidt [1982] have argued that S3 agreement was not marked suffixally in Common Kartvelian.[4] One possibility, entertained by Oniani [1978: 172-178], is that some or all of the Common Kartvelian ancestors of the S3 suffixes were associated with tense, mood and aspect, as they are now, but contained no person and number features. After the separation of Common Georgian-Zan and Prehistoric Svan, these suffixes were 'reshuffled' to give the $S3_{sg}$ and $S3_{pl}$ suffixes for each of the four screeve groups (Table 5).[5] The Svan system of marking $S3_{pl}$ agreement by addition of a suffix (*-x*) to the $S3_{sg}$ form is believed by Schmidt [1982: 67] to be older than the Georgian-Zan system of paired suffixes.

TABLE 5

Hypothesized origin of Georgian-Zan S3 suffixes

CK tense-mood morphemes	*CGZ S3 suffixes [by screeve groups (Table 2)]*
*-a	C. 3sg *-a
*-(e)s	A, B. 3sg *-s; C. 3pl *-es
*-(e)n	D. 3sg *-n; A, B. 3pl *-(e)n
*-(e)d	D. 3sg/pl *-ed

The reshuffling hypothesis is problematic in several respects.[6] First of all, there is no evidence that the forms with and without vowels are descended from a common ancestor. If they are, the reconstructed Common Kartvelian 'tense-mood morphemes' **-(e)s* and **-(e)n* are each associated with several screeve groups in Georgian-Zan, for which no semantic rationale readily presents itself. Secondly, the fate of these morphemes in Svan is unclear: no cognates have been found for most of them. Furthermore, Boeder [1979: 450] points out that Oniani's hypothesis "leaves the absence of these suffixes in the 1st/2nd person unexplained."

The reconstruction of the Common Kartvelian person/number agreement system proposed here agrees in most details with that presented in Oniani [1978] and Schmidt [1982]. Both the Set S and Set O paradigms are structured by the features <±speaker> and <±addressee>. In addition, the feature of number, <±plural>, is present in the Set S system for all feature pairs except <+sp, +ad>, which is redundantly <+pl>; and <-sp, -ad>. The absence of a number opposition in the 3rd person is a common occurrence in the agreement systems of languages throughout the world.

TABLE 6

Reconstructed person/number agreement system for Common Kartvelian

Set S					Set O	
*(x)w-	<+sp, -ad, -pl>	*(x)w-	*-(s_1)t	<+sp, -ad, +pl>	*m-	<+sp, -ad>
		*l-	*-(s_1)t	<+sp, +ad, (+pl)>	*gw-	<+sp,,+ad>
*x-	<-sp, +ad, -pl>	*x-	*-(s_1)t	<-sp, +ad, +pl>	*g-	<-sp, +ad>
*Ø-<-sp, -ad>					*x-	<-sp, -ad>

1.1.3. Constraints upon agreement in Common Kartvelian. Evidence from the attested Kartvelian languages suggests that the agreement morphemes listed above could not cooccur freely with each other. Gamqrelije [1979:48-50] has proposed that the Common Kartvelian verb was constrained by a principle of "prefixal monopersonalism." In the attested languages, usually no more than one person prefix can occur in the verb. When the verb governs two arguments, each of which can control prefixal person marking, hierarchies of agreement set and person determine which of the two will appear in surface structure. The hierarchies are the same in all Kartvelian languages: (a) Set O > Set S; (b) 1st, 2nd > 3rd person.

TABLE 7

Prefixal person marking in early Old Georgian

morphological *subject*	 *object*	*person* *prefix*	 *examples*
2	1	O_1	šen me mo-**m**-c-e "You$_{sg}$ will give it to me"
1	2	O_2	me šen mo-**g**-c-e "I will give it to you$_{sg}$"
3	2	O_2	man šen mo-**g**-c-e-s "S/he will give it to you$_{sg}$"
2	3	S_2?	šen mas mi-**x**-c-e[7] "You$_{sg}$ will give it to him/her"
3	1	O_1	man me mo-**m**-c-e-s "S/he will give it to me"
1	3	$O_3 + S_1$	me mas mi-**x-w**-c-e "I will give it to him/her"

When the morphological subject is 1st person and the morphological object is 3rd person, in Old Georgian and many modern Georgian dialects both S_1 and O_3 appear in the prefixal agreement slot.[8] According to Gamqrelije, this departure from the constraint so rigidly adhered to elsewhere is an innovation dating from the Prehistoric Georgian period [1979: 47-8]. Earlier the S_1 prefix was **xw-* as in Svan, but by the early Old Georgian period it had been reinterpreted as two morphemes, at the expense of the prefixal monopersonalism constraint [see also Klimov 1964: 258].

In the Georgian and Zan verb two agreement morphemes can cooccur in suffixal position. The first morpheme has different forms for 3sg, 3pl and non-3rd person morphological subject, and the second, in word-final position, codes for number. In the following Laz examples, the morphemes in question are represented by *-i/u/es* and *-t* , respectively:

1sg:	b-ğurut'-i "I was dying"	1pl:	b-ğurut'-i-t "we were dying"
2sg:	Ø-ğurut'-i	2pl:	Ø-ğurut'-i-t
3sg:	ğurut'-u	3pl:	ğurut'-es

The situation in Svan is less easy to assess, due to the complex morphophonemic rules which operate in that language. For most screeves, the only suffix sensitive to the category of person is the word-final number agreement marker (1pl/2pl *-(š)d* , 3pl *-x*). The aorist and imperfect screeves, however, are characterized by distinct stems for the S1sg and S2sg forms, versus the S3sg and all plural forms [Topuria 1967: 102-6; 196]. Here are two paradigms in the Lower Bal dialect of Svan:

imperfect

1sg:	xw-a-mara-sgw "I was preparing it"	1pl: xw-a-mara-d "we were preparing it"
2sg:	x-a-mara-sgw	2pl: x-a-mara-d
3sg:	a-mara	3pl: a-mara-x

aorist

1sg:	ot'ex (< a + xw + t'ex) "I returned"	1pl:	ot'äx-d (< a + xw + t'äx + d)
2sg:	at'ex (< a + x + t'ex)	2pl:	at'äx-d
3sg:	at'äx (< a + t'äx)	3pl:	at'äx-x

Schmidt [1982] believes that this harks back to the general pattern of verb inflection in Common Kartvelian: the unmarked S3 form would have served as the stem for the S1pl and S2pl forms, while the

S_{1sg} and S_{2sg} would have employed a second, more marked stem, as diagrammed below:[9]

COMMON KARTVELIAN SET S MARKING

S_{1sg}:	**xw* -stem$_1$	S_{1pl}:	**xw* -stem$_0$-*$(s_1)t$
S_{2sg}:	**x* -stem$_1$	S_{2pl}:	**x* -stem$_0$-*$(s_1)t$
S_3:	stem$_0$ (base form)[10]		

If Schmidt's hypothesis is correct, then the Common Kartvelian verb was characterized by suffixal as well as prefixal monopersonalism: one Set S or O prefix in initial position,[11] and either a number agreement morpheme or a 1sg/2sg stem formant in final position. Although the calculus is a bit involved, there would have been distinct verb forms for each permissible combination of morphological subject and object:

COMMON KARTVELIAN SET S AND SET O MARKING

S_{1sg}/O_2:	**g* -stem$_1$	$S_{1exclpl}/O_2$:	**g* -stem$_0$-*$(s_1)t$
S_{2sg}/O_{1excl}:	**m* -stem$_1$	S_{2pl}/O_{1excl}:	**m* -stem$_0$-*$(s_1)t$
S_{1sg}/O_3:	**xw* -stem$_1$	$S_{1exclpl}/O_3$:	**xw* -stem$_0$-*$(s_1)t$
		S_{1incl}/O_3:	**l* -stem$_0$-*$(s_1)t$
S_{2sg}/O_3:	**x* -stem$_1$	S_{2pl}/O_3:	**x* -stem$_0$-*$(s_1)t$
		S_3/O_{1excl}:	**m* -stem$_0$
		S_3/O_{1incl}:	**gw* -stem$_0$
		S_3/O_2:	**g* -stem$_0$

1.1.4. The Georgian number agreement suffix **-en-.** There is in addition one more agreement morpheme which codes the feature of number. This is the suffix *-(e)n-*[12], which occurs in Old Georgian and some conservative modern dialects (Pshav-Xevsurian, Glola Rachan). It codes agreement in number with plural NOM NPs in the aorist and perfect series of verb screeves.[13] It can occur in any transitive Class A verb, but among Class P verbs, it is restricted to those marked by the preradical vowels *-i-* or *-e-* (this group, which will be referred to as **prefixal Class P verbs,** is composed of prefixal passives and those passives of state and *verba sentiendi* which employ *e* -prefixed forms in the aorist series) [Šanije 1982: 75; Deeters 1930: 61-65; Harris 1985: 303]. The NPs which can control number agreement in *-en-* are either morphological subjects of intransitive verbs, or direct objects of transitive verbs, i.e., absolutive NPs[14] (Table 8):

TABLE 8
Number agreement in *-en-* in Old Georgian
[NPs capable of controlling agreement highlighted]

	CLASS A VERBS		CLASS P VERBS		
			(i/e-prefixed only)		
	Agent	*Patient*	*Recipient/ Experiencer*	*Theme*	*Recipient/ Experiencer*
present series					
agreement	S	O	O	S	O
case	NOM	DAT	DAT	NOM	DAT
aorist series					
agreement	S	***O***	O	***S***	O
case	ERG	***NOM***	DAT	***NOM***	DAT
perfect series					
agreement	O	***S***	——	***S***	O
case	DAT	***NOM***	[+ postposition]	***NOM***	DAT

Here are some early Old Georgian examples (from the 5th-6th century Khanmeti palimpsests edited by Kaǰaia [1984]) illustrating the capacity of a NOM-case NP to control number agreement in *-en-*.

Class A verb
(man) kmn=n=a sasc'aul-n-i
s/he:ERG do:$S3_{sg}$:AbsPL:AOR miracle-PL-NOM
"He worked miracles." [Matthew 21:15]

Cf.:
(man) kmn=a c'q'al-i ğwino-d
s/he:ERG do:$S3_{sg}$:AOR water-NOM wine-ADV
"He made water into wine." [John 4:46]

prefixal Class P verb
še=jrc'un=d=es mcwel-n-i igi da
frighten:INCH:S_{3pl}:AOR guard-PL-NOM the:NOM and

Ø x=i=kmn=n=es vitarca mk'wdar-n-i
(they:NOM) do:PASS:S_{3pl}:AbsPL:AOR like dead-PL-NOM
"The guards became frightened, and they became like dead men." [Matthew 28:4]

In the final example above, note that the *-en-* suffix only appears in the prefixal Class P verb, and not in the *d/n-* suffixed Class P verb *(še=jrc'un=d=es,* not **še=jrc'un=d=n=es).*

Another characteristic of this morpheme is that it only codes number agreement with *syntactically* plural NPs. In Old Georgian, the class of syntactically plural NPs[15] includes NPs headed by nouns marked by the suffixes *-n-* (NOM plural) and *-ta* (DAT/ERG/GEN plural), but excludes those headed by nouns in *-eb-* , although the latter suffix also indicates (semantic) plurality. Šanije [1982:182] cites the following minimal pair from Psalm 73:

{a} šen še=h=musr=en tav-**n**-i
you:ERG crush:S_{2sg}:AbsPL:AOR head-**PL$_n$**-NOM

vešap'-ta-n-i c'q'al-ta šina
dragon -GENPL-PL-NOM water-DATPL in
"You crushed the heads of the dragons in the waters."
[Psalm 73:13]

{b} šen še=h=musr=e tav-**eb**-i
you:ERG crush:S_{2sg}:AOR head-**PL$_{eb}$**-NOM
vešap'-isa-y mis
dragon-GEN-NOM the:GEN
"You crushed the heads of the dragon." [Psalm 73:14]

In {a} the direct object NP *tav-n-i* controls agreement in both its GEN case adjunct *vešap'-ta-n-i* and the aorist-series Class A verb *še=h=musr=en* , while the semantically equivalent *tav-eb-i* controls agreement in neither in {b}.[16]

The distribution of the Georgian morpheme *-(e)n-* is difficult to explain. It is restricted to certain types of finite verbs (Class A and prefixal Class P) in certain screeves (the aorist-series screeves, and those perfect-series screeves which are formed from aorist stems). There are also scattered occurrences of an apparently unmotivated suffix *-(a/e)n*, which may be linked with the agreement morpheme *-(e)n-* in certain stative and medial verb forms. I will present a hypothesis to account for this distribution in the following section. There is no agreement marker attested in Svan or Zan which is functionally comparable to Georgian *-en*. However, as I will attempt to demonstrate below, the evidence suggests that in fact *-en-* was not originally an agreement marker, but rather an indicator of verbal plurality.

2. Verbal plurality in Kartvelian. Verb stems with distinctly plural meaning can be formed in Georgian and Svan. The notion of "verbal plurality" employed in this article is based upon that described by Dressler [1968] in his study of the coding of number in the verb [see also Frajzyngier 1985, Mithun 1988]. The category of plurality in the verb covers different semantic ranges in different languages: iterativity, collective or distributive action, habituality, etc. Whenever the semantic range of verbal plurality includes the numerosity of the arguments, almost invariably the relevant participants correspond to the subjects of intransitive verbs or direct objects of transitives, i.e., "the participants most directly affected by an action" [Mithun 1988: 214; see also Uhlenbeck 1916: 191-195; Dressler 1968: 70-71]. In this paper I will use the term LOCAL ARGUMENT to indicate that argument of the verb — usually a patient or theme — which is especially closely bound to it, as indicated by selectional restrictions and phenomena such as noun incorporation and idioms (for more on this topic see Tuite [1988: §§1.3.3, 2.3.1; 1990]). The local argument normally surfaces as the subject of an intransitive verb or the direct object of a transitive verb.

It is important to distinguish between the indication of a plurality of participants by means of a **verbal plurality marker** from such indication by means of an **agreement marker.** In the former case we have to do with a derivational process. In a study of lexicalization patterns, Talmy [1985: 125-138] noted that number, defined as "the numerosity of the participants — from one to many — behaving as any single argument of a verb," is among the properties likely to be incorporated into the semantic structure of verb roots. Number in this sense corresponds to "an essential aspect of an event, because such numerosity affects how the event is manifested" [Talmy 1985:133]. Agreement, by contrast, is an asymmetrical process by which a feature pertaining to one constituent of the phrase or clause is reflected onto another constituent.[17] It is a syntactic procedure, not a lexical one. A comparison of the two types of verbal number marking in a variety of languages gives the following picture:

[1.] *Number (numerosity) coding associated with the category of verbal plurality:*
  a. Codes number of local arguments only (i.e., absolutive patterning)
  b. Always reflects referential (rather than formal) characteristics
  c. Marking always possible regardless of person, gender, animacy, etc., of NP referring to relevant participants
  d. Category coded by stem suppletion, reduplication, or segmentable morpheme

[2.] *Number marking associated with agreement:*
  a. Any core grammatical relation can be coded. Nominative, ergative or active-stative patterning possible (nominative pattern predominates)
  b. Can reflect purely formal as well as referential characteristics
  c. May not be possible in some person(s), gender(s), etc., or for NPs with inanimate reference
  d. Coded by segmentable morpheme, portmanteau morpheme; occasionally by suppletion (e.g., *I am, we are)*

2.1. Plural verbs in Georgian. Verb-root suppletion occurs in both Old and Modern Georgian [Šanije 1953: 516-538]. A prominent part is played by the root *-sx-* which appears in several suppletive pairs.

Here are some Old Georgian examples [from Šanije 1982: 145-147]. The roots in {a} are transitive, and those in {b} are intransitive.

{a} *-sw-/-sx-* "set, lay"

da=v=sw=a	igi	saq'dar-ta	šen-ta
set$_{sg}$:S$_{1sg}$:AOR	her/him:NOM	throne-DATPL	your-DATPL

"I will set him upon your throne." [Psalm 131:11]

da=v=sx=n=e	mt'er-n-i	šen-n-i
set$_{pl}$:S$_{1sg}$:AbsPL:AOR	enemy-PL-NOM	your-PL-NOM

kweše	perq-ta	šen-ta
under	foot-DATPL	your-DATPL

"I will place your enemies beneath your feet." [Matthew 22:44]

{b} *-v-/-sx-n-* "have [animate theme]" (indirect class P)

g=i=v=i=s	jma-y
have$_{sg}$:O$_2$:S$_{3sg}$:PRS	brother-NOM

"Do you have a brother?" [Genesis 43:7]

m=i=sx=en=an	xut-n-i	jma-n-i
have$_{pl}$:O$_1$:S$_{3pl}$:PRS	five-PL-NOM	brother-PL-NOM

"I have five brothers." [Luke 16:28]

Number suppletivism involving the root *-šx-* (cognate to the Georgian *-sx-* in the above sentences) is also attested in Mingrelian [Klimov 1964:178-9]. As these examples illustrate, Georgian verb-root suppletion reflects features of the local argument, but never features of the transitive agent [see also Harris 1981: 18-21].

There is also a derivational process in Georgian which is sensitive to the number of the local argument. According to Šanije [1953: 262-5] the preverb *da-* (basic meaning "downward") is frequently substituted for other preverbs (especially *ga-* "out," the most commonly-used preverb and the one with most attenuated directional meaning) to denote what he calls "intensity of action" *(mokmedebis int'ensioba).* By this he means multiple, usually serial, completion of the action denoted by the verb [Šanije 1953: 262]. More specifically, for a large number of verbs — over 150, according to Vogt [1971:175] — use of the *da-* preverb implies plurality of the local argument. Some examples [from Šanije 1953: 263-4; Schmidt 1957: 14]:

transitive verbs

man	k'ak'al-i	ga=t'ex=a
s/he:ERG	nut-NOM	break:S_{3sg}:AOR

"S/he cracked the nut."

man	k'ak'l-eb-i	da=t'ex=a
s/he:ERG	nut-PL-NOM	break_{pl}:S_{3sg}:AOR

"S/he cracked the nuts."

man	bavšv-i	ga=zard=a
	child-NOM	grow:S_{3sg}:AOR

"S/he raised the child."

man	bavšv-eb-i	da=zard=a
s/he:ERG	child-PL-NOM	grow_{pl}:S_{3sg}:AOR

"S/he raised the children."

intransitive verbs

k'vercx-i	ga=laq'=d=a
egg-NOM	addle:S_{3sg}:AOR

"The egg went bad."

k'vercx-eb-i	da=laq'=d=a
egg-NOM	addle:S_{3sg}:AOR

"The eggs went bad."

bavšv-i	ga=i=zard-a
child-NOM	grow:PASS:S_{3sg}:AOR

"The child grew up."

bavšv-eb-i	da=i=zard=nen
child-PL-NOM	grow_{pl}:PASS:S_{3pl}:AOR

"The children grew up."

According to Schmidt [Schmidt 1957: 14] the *da-* preverb can also signal "adverbial plurality," as in the following instances: *da=k'ocn=i=s* "somebody kisses one person many times, or several people once," *da=srial=eb=s* "somebody slides or skates back and forth," *da=q'id=i=s* "somebody sells many things, or goes around selling things." Šanije [1953:263] contrasts *p'ur-is ga=č'ra* (bread-GEN cutting:NOM) "cutting bread in two," with *p'ur-is da=č'ra* (bread-GEN cutting$_{pl}$:NOM) "cutting bread into many slices."[18]

2.2. Plural verbs in Svan. A second verb-stem derivational process connected with number is found in some dialects of Svan. The suffix *-a:l-* and its variants *(-ie:l-, -ə:r)* can signal plurality of the local argument, or of the action in general [Deeters 1930: 66-8; Šarajenije 1954; Schmidt 1957: 19-20; Topuria 1967: 233]. The following Upper Bal examples are taken from Šarajenije [1954: 192, 195]:

Plural direct object

{a} xam-är ady=a:l=e
pig-PL:NOM drive.away$_{pl}$:S3:AOR
"He drove the pigs away."

{b} o=xw=šk'ad=a:l=a lucman-är
forge$_{pl}$:O3:S3:PERF nail-PL:NOM
"He has forged the nails."

Plural intransitive subject

{c} mäg ušxwid än=ərd=ä:l=ä:n=x
all together grow$_{pl}$:PASS:S3:AOR
"They all grew up together."

Here are some pairs of verbal nouns from Gujejiani & Palmaitis [1985]:

Svan verbal nouns (Upper Bal dialect)

li=šxbi "to sew one thing"	li=šxbiy=e:l=i "to sew many things"
li=je:lwe "to sweep sthg"	li=je:lw=ä:l=i "to sweep sthg many times"

li=c'b=i:ne "to dangle (tr/intr)	li=c'e:b=ä:l "to dangle many things, to dangle on sthg many times"

In addition, many Class A denominal verbs denoting activities (events perceived in terms of their temporal duration rather than change of state) are derived with the suffix *-a:l-* [Č'umburije 1981:149ff], e.g.:

Activity verbs in -a:l- (Upper Bal)
li=bəlkow=ä:l=i "to play cards" [< *bəlkow* , a name of a card game]
li=wğw=in=ä:l=i "to work with a yoked ox" [< *uğwa* "yoke"]
li=lc=e:r=ä:l=i "to water, irrigate" [< *lic* "water"]

All of these denominal verbs in *-a:l-* are characterized by atelic aspect, in which respect they resemble the Georgian medial (or medioactive) verbs [Nozaje 1974; Holisky 1981]. Overall, the semantic range of the Svan verbal pluralizer *-a:l-* is quite similar to that of the Georgian preverb *da-* discussed previously.

2.3. Common Georgian-Zan *-en- *as a verbal plurality morpheme.* I propose that at some earlier stage, perhaps in the Common Georgian-Zan period, the morpheme **en* — the ancestor of the number agreement morpheme described in §1.1.4 — was a derivational suffix with a semantic range similar to that of the Svan pluralizers discussed in the previous section. I have two primary reasons for believing so, which will be presented here:

(A.) *The position of -en- in the verb.* Although it is not completely agglutinative, the Kartvelian verb has essentially the same sequence of morphemes in all three languages. Its morphemic composition can be diagrammed as follows [cf. Deeters 1930: 6-7]:

Morphemic composition of the Kartvelian verb

$$[\text{preverb}]_0 +$$
$$[\text{S/O}_1\text{=}[\text{ver}_2\text{=}[[\text{root}_3]_a\text{=inch/caus}_4\text{=psf}_5\text{=plural}_6]_b\text{=impf}_7\text{=tns/md}_8]_c\text{=S}_9\text{=num}_{10}]_d$$

Structural levels:

a. Verb root.
b. Components occurring in nonfinite as well as finite verb forms: root, causative formant (slot 4), present-stem formant (slot 5), pluralizer (slot 6). All of these components occur in Svan verbal (tr/intr) finite verbs. There is also an inchoative-intransitive formant *(-n/d-),* for which no comparable morpheme has been found in Svan [Topuria 1967: 40; but cf. Klimov 1964: 79]; it may represent a Common Georgian-Zan innovation.
c. Components indicating verb class and screeve: the above plus the version vowel (slot 2), imperfect-stem formant (slot 7) and tense/mood vowel (slot 8).
d. The fully-inflected finite verb: all of the above with the addition of the Set S/O person agreement prefix (slot 1), the Set S person agreement suffix (slot 9) and the number agreement suffix (slot 10).

Examples:

OLD GEORGIAN:
c’ar$_0$+Ø$_1$=a$_2$=vlin$_3$=**n**$_6$=od$_7$=i$_8$=s$_9$
send.away:S$_{3sg}$:O$_3$:AbsPL:MxCNJ
“he would send them” [Mark 3:14 (Adish gospels, 897)]

SVAN:
ži$_0$ + x$_1$=a$_2$=nd$_3$=un$_4$=**a:l**$_6$=wn$_{(4)}$=e$_8$=x$_{10}$[19]
want$_{pl}$:CAUS:S$_{3pl}$:O$_3$:FUT
“they will cause him to want them” [Topuria 1967: 234 (Lashx dialect)]

In all three Kartvelian languages the Set S and O person agreement morphemes are positioned at the extreme periphery of the verb. The Svan and Georgian pluralizers, by contrast, occupy a position much closer to the verb root, along with other stem-derivational morphemes. Although Georgian *-en-* is only attested as a number agreement marker, its location in the verb betrays its original function.

(B.) *The distribution of -en- and cognate suffixes in Georgian and Zan.*

(B^{1}.) en-*agreement in Georgian.* As was noted in the preceding section, this morpheme is restricted to certain types of Class A and prefixal Class P verbs in the aorist-series screeves and to those perfect-series screeves which are formed from aorist stems. There does not appear to be any semantic explanation ready to hand which would cover these and only these verb forms. One feature which is shared by most of the verbs which can be marked by *-en-* in Old Georgian is that they are "highly transitive" in the sense developed by Hopper and Thompson [1980]. Specifically, they are characterized by punctual aspect (item D on Hopper and Thompson's list of transitivity parameters [1980: 252]) and a transitive semantic structure (item A). The former criterion accounts for the restriction of *en*-agreement to aorist stems: in early Old Georgian the aorist stem was associated with punctilear aspect, and was opposed to the linear, durative aspect of the present-series stem [Mač'avariani 1974; Schmidt 1984].

Concerning the second point, Gamqrelije [1979:46-47] has noted that prefixal passives (those Class P verbs marked by the pre-radical vowels *-i-* and *-e-* in both the present and aorist series) are semantically more like true passives than the other types of Class P verbs. Many prefixal passives give evidence of having transitive semantic structures. For example, they allow oblique agent phrases, while root and suffixal Class P verbs almost never do [Harris 1985:60-61]. Oblique agent phrases (marked by the postpositions *mier* 'by' or *gan* 'from') are not especially common in Old Georgian texts, but when they do occur it is with a prefixal — and never with a root or suffixal — Class P verb, e.g.:

mi=x=i=q'wan=a igi angeloz-ta-gan
take:PASS:[P]:S$_{3sg}$:'O$_{3}$':AOR he:NOM angel-GENpl-from
c'iağ-ta abraham-is-ta.
bosom-DATpl Abraham-GEN-DATpl
'He was taken by the angels to the bosom of Abraham.' [Luke 16:22]

In those cases where both prefixal and root or suffixal Class P stems can be derived from the same verb root the difference is easily discerned. Compare the following two Modern Georgian sentences:

(i) xač'ap'ur-i cxveb=a (*ded-is mier)
cheesebread-NOM bake[P_{root}]:S_{3sg}:PRS mother-GEN by
'The cheesebread is baking (*by mother).'

(ii) xač'ap'ur-i i=cxob=a
cheesebread-NOM bake:PASS:[$P_{i\text{-prf}}$]:S_{3sg}:PRS
ded-is mier
mother-GEN by
'The cheesebread is being baked by mother.'

The contrast between (i) and (ii) is similar to the contrast between what are sometimes termed "unaccusative" (or "ergative") and passive constructions in English [Keyser & Roeper 1984]. The semantic structures underlying the two constructions are quite different:

TRANSITIVE SEMANTIC STRUCTURE
(a^1) *transitive construction:*
agent ⇒ subject patient/theme ⇒ direct object

(a^2) *passive construction:*
agent ⇒ oblique NP patient/theme ⇒ subject

INTRANSITIVE SEMANTIC STRUCTURE
(b) *"unaccusative" construction:*
patient/theme ⇒ subject

By the time of the earliest Georgian texts, *en*-agreement was no longer motivated by the above rule. Although it plausible that in the late Prehistoric Georgian period all prefixal passives had transitive semantic structures, by the time of the earliest Georgian texts there were already some exceptions to this pattern (though most prefixal passives then, and in modern Georgian, continue to be underlyingly transitive). Secondly, those Class P verbs which have aorist-series forms marked by the pre-radical vowel *-e-* (passives of state and *verba sentiendi)* are also capable of *en*-agreement in Georgian. Almost all of these latter verbs are stative, and the original motivation for the presence of *-en-* in this case can be linked with that of a morpheme which I believe to be cognate in the verbs to be discussed in the following section.

(B^2.) *The suffix -(a/e)n- in Georgian stative verbs and in Zan.* There are about a half-dozen Georgian stative verbs for which a synchronically-unmotivated suffix*-(a/e)n-* appears in the S3pl present and present-indicative forms; e.g., *c'v=an=an* "they are lying down" (cp. 3sg *c'ev=s)*; *sx=en=an* "they are sitting." For some of these verbs this suffix also appears in the imperfect stem: *x=gv=an=d=a* "somebody resembled somebody" (present *x=gav=s);* *x=a=kw=n=d=a* "somebody had sthg" (present *x=a=kv=s).* All of these belong to the group of root Class P verbs, which is believed to be among the most archaic verb types [Harris 1985: 59].

Šanije [1953: 426-427] sees in the *-(a/e)n-* morpheme the descendent of a present-stem formant (slot 5) which has lost its original function. More recently, Harris [1985: 222-228] has argued that this suffix, and also the *-n-* element of the S3pl suffix used in the present and present-perfect screeves of all Class P verbs in Zan,[20] is to be equated with the Georgian pluralizer *-en-*. This proposal, I believe, is fundamentally correct, although my interpretation of the data differs from hers in some respects.

If one accepts the hypothesis that Common Georgian-Zan **en* was a verbal plurality morpheme along the lines of Svan *a:l* , then one can make predictions about where its traces would most likely be found in the daughter languages. First of all, where it is attested in connection with plural number agreement, it should only correlate with local arguments (absolutive NPs). The Zan S3sg suffix *-n-an,* as Harris notes, is only used by Class P verbs, all of which are intransitive. It never appears in correlation with transitive subjects. In the other environments where the ancestor form of Zan *-n-* would be expected to appear — in correlation with plural direct objects and non-3rd person intransitive subjects — it is no longer in evidence. A second environment where remnants of Common Georgian-Zan **en* would be likely to turn up is in verbs connoting iterative, continuative or habitual *Aktionsart,* these being the semantic features most often associated with verbal plurality marking [Dressler 1968: 60-65, 74-77]. Both Zan *-n-* and the anomalous Georgian formant*-(a/e)n-* occur in the present series, which was associated with durative aspect in Old Georgian [Mač'avariani 1974; Schmidt 1984]. In the case of Georgian, one notes that *-(a/e)n-* only occurs in stative verbs. Such verbs denote situations, qualities, etc. which of necessity extend over a period of time. For this reason they do not have opposed present-series and aorist-series forms, and the conflation of the present and present-iterative screeves observed in many Old Georgian texts is particularly common for stative verbs [examples in Sarǰvelaje 1984: 425ff]. A similar situation may have obtained at an earlier time in regard to verbal-plurality marking, i.e., there was no semantic opposition between iterative and non-iterative *Aktionsart* for stative verbs, and therefore the appearence of the **en* formant was not motivated as it was for other verb types. Presumably it became a purely lexical component of these stative verbs at an early stage.

(B^{3}.) *The suffix -n- in Georgian medial verbs.* Finally, what may be traces of Common Georgian-Zan **en* can be observed in the declension of Georgian medial (medioactive) verbs. Although usage varies, many Georgian speakers employ a stem formant *-n-* in the perfect-series screeves of medial verbs [Šanije 1953: 491]; e.g., *Ø=u=duğ=n̲=i=a* "it has boiled" (aorist *i=duğ=a*); *Ø=e=tamaš=n̲=a* "s/he had played." In Šanije's opinion, this suffix originated in the pluperfect, and later spread to the present perfect [1953: 455]. In the few Early Old Georgian examples, the formant *-n-* is found only in the pluperfect or other perfect-series screeves formed from the aorist stem, e.g., *x=e=q'iv=n̲=o=s* "<the rooster> will have crowed" (perfect conjunctive) [John 13:38 (Khanmeti gospels, 5th-6th c.)]. If this otherwise-unexplained suffix is in fact descended from the pluralizer **en* , then it would have had a semantic range very close to that of the Svan verbal-plurality formant *-a:l-* which appears in many medial verbs (examples in section 2.2. above).

It is clear from the above data that at some time in the Prehistoric Georgian period the range of usage of **en* came to be restricted in a way not characteristic of its cognate in Zan, nor of Svan*-a:l-*. For all non-stative verb types **en* was restricted to verb forms based on the aorist stems — this is the case in Old Georgian for both the number agreement marker *-en-* and the formant *-n-* attested in some medial verbs. More precisely, it was in complementary distribution with the present-stem formants (slot 5 in the verb's morphological structure). This is true even of the stative verbs described in B^2 above (hence Šanije's hypothesis that this *-(a/e)n-* morpheme was itself once a present-stem formant).

The conclusion I wish to derive from these data is that the various Georgian and Zan *-(a/e)n-* suffixes discussed here reflect two different stages in the history of the Common Georgian-Zan pluralizer **en:*

STAGE I [Common Georgian-Zan period] — The formant**en* is used to derive plural verb stems. The semantic range of this suffix is comparable to that of Svan-*a:l-:* plurality of local argument; iterative, habitual, continuative Aktionsart. Remnants of this stage include the *-n-* in the S_{3pl} suffix of certain intransitive verbs in Zan, and the *-(a/e)n-* formants appearing in some screeves of Georgian stative and medial verbs (which are, of necessity, aspectually continuative).

STAGE II [Prehistoric Georgian period] — The pluralizer **en* came to be restricted to verb forms based on aorist stems. Later it was reanalyzed as an inflectional suffix, marking number agreement with the direct objects of verbs with transitive deep structures. The number agreement suffix *-en-* of Old Georgian and some modern dialects is a continuation of this morpheme.

3. Verbal plurality and nominal plurality. A survey of verbal *Aktionsart* categories in a large number of languages led Dressler [1968] to conclude that the varieties of nominal number/numerosity marking ('true' plurals, collectives, distributives, etc.) have their semantic counterparts in the observed varieties of verbal plurality. He proposes that the semantic features subsumed under plurality be conceived of as supracategorial, taking their more specific meanings depending on which lexical category they are applied to.[21]

In some languages, the same formal means are used to express a particular type of numerosity in the noun and also the corresponding category of plural *Aktionsart* in the verb. Mithun [1988] reports several instances of this from North American languages: In Nass Tsimshian, reduplication indicates distributive plurality (of events and/or participants) in the verb, and distributive plurality in the noun; a similar use of reduplication is observed in Southern Paiute, Takelma, Tonkawa and Barbareño Chumash. The Maidu suffix *-to-* indicates iteration or plurality of the direct object in verbs, and distributive plurality in nouns [ibid: 226-227]. Frajzyngier [1985] reports examples of similar phenomena from a variety of language families.

In this section I will examine the nominal and verbal pluralizers attested in the Kartvelian languages. The evidence points to the existence of a well-developed category of "number" that was marked by morphemes shared by noun and verb in Common Kartvelian.

3.1. Svan -a:l- *and Geo* -en-. Both of these verbal pluralizers show a correlation with local arguments and bear a strong resemblance to nominal pluralizers in their respective languages. They are juxtaposed in Table 9:

TABLE 9

Comparison of Georgian *-en-* and Svan *-a:l-*

verbal morpheme	distribution	related forms
Georgian *-(e)n-*	Correlated with *n* -plural NOM NPs in aorist series and those perfect series forms based on aorist-series stems; Used in Class A and *i/e* -prefixed Class P verbs.	*-n-* [NOM plural]

Svan *-a:l- / -ie:l- / -ə:r-*	1. Code local-argument plurality 2. Code iterative action; 3. Appear in many medial verbs; Used in both finite and nonfinite verb forms	*-ar* [22]/ *-e:r / -i:r / -a:l / -a:ru / -ie:l / -ə:r/*... [nominal pluralizers]

In his analysis of Kartvelian *Aktionsart*, Schmidt [1957: 19-21] conjectures that the verbal and nominal pluralizers juxtaposed above are etymologically related.[23] For Svan, Sharadzenidze [1954: 202] reconstructs six pluralizers *(*a:r, *u:r, *a:l, *a:d, *u, *a)*, the first three of which are common to verbs and nouns. In terms of their semantic range, Svan *-a:l-* , *-i-e:l- (<*a:l)* and-*ə:r- (<*u:r)* can be conceived of as supracategorial pluralizers of the type proposed by Dressler [1968], for which plural reference and plural *Aktionsart* can be regarded as "combinatorial variants of the sememe 'plurality' for nouns and verbs, respectively" (see note 20).

The Georgian pluralizer *-en-* is clearly a number agreement morpheme, but the evidence summarized in the preceding section suggests that it is descended from a Common Georgian-Zan verbal plurality formant. Its nominal counterpart only appears in the NOM and VOC cases of one of the two plural declensions employed in Georgian (Table 10):

TABLE 10

Georgian plural declension paradigms

NOMinative	kal-**n**-i	kal-eb-i "women"
ERGative	kal-ta	kal-eb-ma
DATive	kal-ta	kal-eb-s
GENitive	kal-ta	kal-eb-is
INStrumental	—	kal-eb-it
ADVerbial	—	kal-eb-ad
VOCative	kal-**n**-o	kal-eb-o

The VOC is believed to be a relatively recent arrival to the Georgian declensional system [Vogt 1947: 99], and its form is derived from the NOM. The fundamental opposition, then, is NOM (rectus) *-n-* :: DAT/ERG/GEN (oblique) *-ta*. Two-stem (rectus/oblique) declension was evidently also characteristic of Common Kartvelian, as it is widespread in Svan nominal morphology, and is found in the declension of 3rd person pronominals in Georgian and Zan.

Alongside their obvious differences, Georgian *-n-* and the various Svan nominal pluralizers share one characteristic: both are treated as markers of formal plurality by the syntax. In this respect they differ sharply from Georgian *-eb-* and its Zan cognate *-em-* / *-ep-* . Nouns marked with the latter can control plural number agreement only if they refer to animate beings, if even then.

number agreement with inanimate subjects

-ar- [Svan]

lekwr-iš	bač-är	eser-i	qečn=i=**x**
mill-GEN	stone-PL:NOM	QUOT-also	wear.out:S3$_{pl}$:PRS

"Millstones wear out also." [Davitiani 1973: 56]

-n- [Old Georgian]

rabam-n-i	kwa-n-i	ar=i=**an**	ese
how.many-PL-NOM	stone-PL-NOM	be:S3$_{pl}$:PRS	this:NOM

"How many stones there are!" [Mk 13:1 (Adish)]

no number agreement with inanimate subjects:

-eb- [Old Georgian]

rabam-i	kw-eb-i	ar=s
how.many-NOM	stone-PL-NOM	be:S3$_{sg}$:PRS

"How many stones there are (lit. 'is')!" [Mk 13:1 (Khanmeti)]

-ep- [Mingrelian]

xal-ep-i	ko=čan=s
tree-PL-NOM	stand:S3$_{sg}$:PRS

"The trees are standing (lit. 'is standing')." [Qipshidze 1914: 5]

3.2. Georgian/Zan -eb- / -em- / -ep-. The above facts indicate that the ancestor of Georgian *-eb-* and Zan *-em-* / *-ep-* was something other than a 'true' plural marker. A number of authors have, in fact, suggested that these suffixes are descended from a *collective* marker, denoting a group taken as a whole rather than as a plurality of individuals. According to Blake [1932: 252], in early Old Georgian texts *-eb-* most often appeared with "words which denote concrete objects ... when an indefinite quantity is conceived of and not a specific number." In his grammar of Old Georgian, Zorell [1930: 19] glosses *švil-i* as 'Kind,' *švil-n-i* as 'Kinder,' but *švil-eb-i* as 'etwa = eine Kinderschar.' There are also several derivational morphemes in Georgian which may be related to *-eb-* : the suffixes *-ev-* (which occurs in some toponyms, e.g., *Vašl-ev-i* "Apples" [Šanije 1953: 138]) and *-ob-* (used to form abstract nouns and collectives, e.g., *st'udent'-ob-a* "student life; student body," *st'umr-ob-a* "being a guest; the guests [as a group]"). The failure of number agreement to occur with *eb-* suffixed nouns in Old Georgian can be explained as stemming from the time when *-eb-* was used to derive formally singular collective nouns.[24]

Harris [1984; 1985: 194ff] accepts this interpretation, and extends it a step further. In her view, the Common Georgian-Zan ancestor of-*eb-* was a morpheme which indicated collectivity in nouns, and durative aspect in verbs. Its descendents in the verbal system are the present-stem formants — suffixes added to the basic stem in all present-series screeves. Most of the Georgian and Zan present-stem formants are composed of a vowel followed by a labial consonant; e.g., Georgian *-eb-, -ob-, -av-, -am-, -em-;* Zan *-um-, -im-, -ap-, -ep-, -em-*. Harris points out several examples of "haphazard" correspondences of labial phonemes between cognate forms in Georgian-Zan [1985: 208 note 3], and on this basis argues that most of the present-stem formants are derived from a single Common Georgian-Zan antecedent **ev*.[25] Its function in verbs was to derive the durative-aspect present-series stem from the more basic aorist-series stem. Added to nouns, it formed the collective.[26]

3.3. Common Kartvelian verbal/nominal plurality. There is good evidence that at least three Kartvelian plural morphemes were at one time productively used by verbs and nouns alike: Common Georgian-Zan **(e)n,* Common Georgian-Zan **ev* (and/or **eb)* and the Svan pluralizers **a:l / *a:r / *u:r*. In each case the semantics of the attested nominal and verbal uses are consistent with Dressler's crosslinguistic findings. In view of their functional parallelism, one might wonder if Common Georgian-Zan **(e)n* is cognate with any of the Svan pluralizers. The evidence from Kartvelian comparative phonology is suggestive but not convincing [Oniani 1978:210-211]. The alternation [n]:[l] is observed both within Svan (e.g., Upper Bal ***l**ic*, Becho ***n**ic* "water" [Topuria 1967: 168]) and within Georgian (e.g., Std Georgian *c'e**l**-i*, Kaxetian *c'e**n**-i* "year" [Gamkrelidze and Machavariani 1982: 45]). There is, however, no systematic [n]:[l] correspondence between the two languages [Schmidt 1962: 86-87]. The correspondence between the vowels (CGZ [e]: Svan [a:]) is also problematic.

Although it is highly unlikely that all of the pluralizers attested in the daughter languages are descended from a single pair of antecedents, it is my belief that they fall into two semantic groups. If my hypothesis is accurate, Common Georgian-Zan, and perhaps Common Kartvelian, was characterized by a three-way plurality opposition:[27]

	nominal	**verbal**
Ø:	(unmarked)	(unmarked)
**en* / **a:l* (others?):	'true' plural	distributive-iterative
**ev* / **eb* (others?):	collective	durative

In principle, nominal and verb stems came in threes, though there were, no doubt, exceptions. Certain stems had specifically plural meaning (examples in §2.1). Verbs which of necessity denoted states or activities extending over a period of time, rather than focusing on a change of state (i.e., stative and medial verbs) may have had defective stem paradigms. One certainly sees evidence of this in Old Georgian: in general, neither statives nor medials had opposed present-stem (durative) and aorist-stem (punctiliar) forms [Šanije 1953: 477, 498].

In regard to nominals, it may have been the case that only certain Common Kartvelian nouns were regularly marked for number. The number of other types of nouns was inferred from context, quantifiers, or the presence of a plural verb. This manner of coding number is characteristic of many languages of North America

[Mithun 1988]. Instances of formally singular nouns in conjunction with a plural verb are found here and there in Georgian texts, especially when the noun in question refers to an object most often met with in sets or groups [Šanije 1953: 39; Imnaišvili 1957:283-5]; e.g.:

> purcel-i misi ara da=s=cwiv=i=s
> leaf-NOM its:NOM not fall$_{pl}$:S3$_{sg}$:PERM
> "Its leaves will not fall off." [Psalm 1:3]

> c'am Ø=u=q'w=n=a mas simon-p'et're
> eyelash:NOM do:S3$_{sg}$:O3:AbsPL:AOR him:DAT S-P:ERG
> "Simon Peter winked (lit. did his eyelashes) at him." [John
> 13:24 (T'beti)]

The referential plurality of *purcel-i* "leaf" and *c'am* "eyelash" is coded only by the plural verb roots: *-cwiv-* "many fall" contrasts with *-vrd-* "one falls," and the stem-*q'w=n-* "do many" includes the *en* pluralizer.

4. Conclusion. Agreement, as Nichols [1985] defines it, is an asymmetric process: one constituent merely reflects a feature proper to another constituent. The system reconstructed here was characterized by considerable symmetry between noun and verb. If both constituents coincided in bearing one of the above morphemes, it was because each was independently marked for the category. In comparison to nominal and verbal plurality marking, true number agreement was weakly developed in Common Kartvelian, being confined to the 1st and 2nd person of the Set S affix system. When a rich number-agreement system evolved in Kartvelian, it was — especially in Georgian-Zan — accompanied by the breakdown of the old plurality system. Morphemes which had been the common property of noun and verb morphology split into distinct nominal and verbal desinences, which went off in their separate directions. Some of the morphemes which had once been verbal-plurality markers were reanalyzed as agreement morphemes (e.g., Old Georgian *en,* perhaps Georgian dialectal *q'e);* they became markers of reflected, rather than proper, features. Others took on entirely new functions (e.g., Georgian-Zan _eb/ep_ as present-stem formants; Zan _-n-_ in S$_{3pl}$ of Class P verbs), and some presumably were lost. To sum up: the Kartvelian morphological system has always coded number in both noun and verb, but the nature of that coding appears to have changed profoundly.

ABBREVIATIONS

Case: NOMinative, ERGative, DATive, GENitive, INStrumental, ADVerbial.
Verb forms:
present series: PRS (present), IMP (imperfect), CNJ (conjunctive), FUT (future).
aorist series: AOR (aorist), PERM (permansive), MxCNJ (mixed conjunctive)
perfect series: PERF(present perfect)
Verb class: [A], [P].
Other: PASSive, INCHoative, AbsPL (plural absolutive); QUOTation particle.

Hyphens (-) indicate that the segmentation of the gloss mirrors the morpheme composition. The double hyphen (=) or colon (:) indicates that no such 1-to-1 glossing of morphemes is attempted. In the verb gloss, the person marker agreeing with the (real) subject is given first.

APPENDIX

This article has focused upon the Kartvelian pluralizers for which reasonably plausible reconstructions are possible. In this appendix I will touch upon some of the remaining plural morphemes which are attested in one or the other of the Kartvelian languages, and review some attempts to reconstruct their earlier distribution.

A. Georgian -t(a): This is the oblique-plural desinence paired with NOM-plural _-n-_ (see Table 10). The attempt has been made [Čikobava 1954] to motivate a connection between this morpheme and the S$_{1/2}$ number agreement suffix (CK _*(s$_1$)t)_. Čikobava proposes that _-t(a)_ is "older" than _-n-)_ and was used to mark the plural of nouns in all contexts, e.g.: _*k'ac-t v-ar-t_ 'man-PL S$_1$-be-PL' = "we are men" [ibid: 74].

More likely to be cognate with _-t(a)_ is the Georgian suffix _-et_ , which used to derive place names (e.g., _somx-et-i_ "Armenia" < _somex-i_ "Armenian [person]") [Vogt 1947]. Related forms are found in Zan (e.g., _Zan-at-i,_ the name of a village in Mingrelia [Klimov 1964: 80]) and Svan (_qae-yšd_ < _*qev-it-i,_ a village name from Upper Svaneti [Č'k'adua 1987: 207]). This suffix was also added to names in Old Georgian to derive nouns denoting a group centered in some sense around an individual, e.g., _p'avle-et-n-i_ 'Paul-et-PL-NOM' = "Paul and his companions," *vaxt'ang-et-i* “the royal house founded by King Vaxtang” [Imnaišvili 1957: 385ff; Boeder 1987]. In this latter context, *-et* codes the sort of notional number associated with the plural forms of 1st and 2nd person pronouns: ‘we’ = ‘I and a group centered around me,’ ‘you_{pl}’ = ‘you_{sg} and a group centered around you_{sg}.’ Even so, an etymological link between *-et* and the $S_{1/2}$ number agreement suffix *(s1)t seems far-fetched.

B. Svan -x . The origins of this suffix, which codes number agreement for the 3rd person in Set S and the 2nd and 3rd persons in Set O, remain obscure. Marr and some of his disciples linked it with a superficially similar Abkhaz plural morpheme. At the end of her article on Svan pluralizers, Sharadzenidze [1954: 203] avers that the *-x-* element in the word *yerxi* “some [people]” may have plural meaning (cp. *yer* “somebody, something”), but draws no further conclusions. Jijiguri [1935] proposed an etymological link with the particle *q'e* (see below), although there is no sound correspondence between Svan and Georgian which would support this hypothesis [Schmidt 1962: 70-71].

C. Georgian q'e . This postclitic is used in some contemporary Georgian dialects, and sporadically attested in medieval texts [Deeters 1930: 60-61; Čikobava 1968: 276ff; Tuite 1988, 1989]. Two distinct functions have been reported for *q'e* and its phonological variant *k'e:* (1) In several eastern and northwestern dialects, it codes the plurality of 2nd and 3rd person NPs when they function as sentential topics, and would otherwise not be able to control number agreement (i.e., because they are crossreferenced by Set O agreement affixes). (2) In some areas of north-central and eastern Georgia, *q'e* indicates habitual or permansive Aktionsart. Only in the far eastern provinces (Saingilo, Kiziqi, Fereidan) is *q'e* used with both functions.

It is highly unlikely that the contexts in which this morpheme appears as a number agreement marker represent its original distribution. Compare the agreement system shown below with the early Old Georgian pattern in Table 3. In all modern Georgian dialects, the Set O paradigm lost the inclusive/exclusive opposition and was realigned according to the feature matrix of Set S. In most regions, either the $S_{1/2}$ number agreement marker *-t* or the clitic *q'e* was used to fill in gaps in the paradigm where no means of coding number was inherited from Prehistoric Georgian.

TABLE 11

Person/number agreement system in east and northwest Georgian dialects

Set S					Set O				
v-	<1, -pl>	v-	-t	<1, +pl>	m-	<1, -pl>	gw-		<1, +pl>
Ø-	<2, -pl>	Ø-	-t	<2, +pl>	g-	<2, -pl>	g-	**-q'e**	<2, +pl>
-s/a/n	<3, -pl>		-en/es	<3, +pl>	h/Ø-	<3, -pl>	h/Ø-	**-q'e**	<3, +pl>

If one were to hypothesize that *q'e* was once a verbal-plurality morpheme similar to those of modern Svan, one would also have to accept that it was only correlated with the plurality of local arguments at that time. Later in the Prehistoric Georgian it would have been reanalyzed and employed in a new range of contexts, just as the former $S_{1/2}$ number agreement marker *-t* was in modern literary Georgian and some other dialects. This would have occurred in connection with a shift in the morphosyntactic orientation of the dialects in question [Tuite 1988 §9].

ACKNOWLEDGMENTS

This paper was written while the author was a post-doctoral research fellow at the Institute of Oriental Studies of the University of Tokyo. I wish to express my gratitude to the Institute and to the Japan Society for the Promotion of Science, which has administered the fellowship program. I also acknowledge the many useful and thought-provoking comments supplied by an anonymous referee for this publication.

BIBLIOGRAPHY

Andğuladze, Nodar. 1968. *K'lasovani da p'irovani uğvlilebis ist'oriis zogi sak'itxi iberiul-k'avk'asiur enebši* [Some questions concerning the history of class and person conjugation in the Ibero-Caucasian languages]. Tbilisi: Mecniereba.

Arabuli, Avtandil. 1984. *Mesame seriis nak'vteulta c'armoeba da mnišneloba jvel kartulši* [The form and meaning of series III screeves in Old Georgian]. Tbilisi: Mecniereba.

Aronson, Howard. 1979. Towards a typology of transitivity: the strange case of the Georgian subject. *CLS parasession on the elements*, pp. 297-306. Chicago Linguistics Society.

_____. 1984. Homonymy in the Georgian verbal system. *Folia Slavica* 7: 21-37.

Blake, R. P. 1932. Khanmeti palimpsest fragments of the Old Georgian version of Jeremiah. *Harvard Theological Review* 25: 225-272.

Boeder, W. 1979. Ergative syntax and morphology in language change: the South Caucasian languages. In F. Plank (ed.) *Ergativity: towards a theory of grammatical relations*, pp.435-480. Orlando: Academic Press.

_____. 1987. Anmerkungen zum Pluralsuffix *et* im Altgeorgischen. In press: to appear in *Folia Slavica* .

Čikobava, Arnold. 1936. *Č'anuris gramat'ik'uli analizi t'ekst'ebiturt.* [A grammatical analysis of Laz, with texts]. Tbilisi: Mecniereba.

_____. 1940. Mesame p'iris subiekt'is ujvelesi nišani kartvelur enebši [The oldest third-person subject marker in the Kartvelian languages]. *Ak'ad. N. Maris saxelobis enis, ist'oriisa da mat'erialuri k'ult'uris inst'it'ut'is moambe* V-VI: 13-46.

_____. 1954. Mravalobitis supiksta genezisatvis kartulši [On the genesis of the Georgian plural suffixes]. *Iberiul-k'avk'asiuri enatmecniereba* 6: 67-76.

_____. 1968. *Mart'ivi c'inadadebis p'roblema kartulši, I: kvemdebare-damat'ebis sak'itxi jvels kartulši* [The problem of the simple sentence in Georgian, I: the issue of subject and object in Old Georgian]. Tbilisi: Mecniereba.

Č'k'adua, Ambako. 1987. T'op'onimik'ur šesat'q'visobata jiritadi p'rincip'ebi [Basic principles of toponymic correspondences], *Onomastika* I: 181-228.

Č'umburije, Zurab. 1981. Nasaxelari zmnebi svanurši. [Denominal verbs in Svan]. In *Ak'ak'i Šanijes* [To Akaki Šanije] ed. by Š. Jijiguri, pp. 145-151. Tbilisi: Tbilisi University Press.

Comrie, Bernard. 1975. Polite plurals and predicate agreement. *Language* 51: 406-418.

Davitiani, Aleksi. 1973. *Svanuri andazebi* [Svan proverbs]. Tbilisi: Mecniereba.

Deeters, Gerhard. 1930. *Das kharthwelische Verbum: vergleichende Darstellung des Verbalbaus der südkaukasischen Sprachen.* Leipzig: Markert und Petters.

Dressler, Wolfgang. 1968. Studien zur verbalen Pluralität. *Sitzungsberichte der philosophisch-historischen Classe der Österreichischen Akademie der Wissenschaften* 259, Abh 1.

Frajzyngier, Z. 1985. Ergativity, number and agreement. *Berkeley Linguistics Society* 11: 96-106.

Gamkrelidze, T. V. and G. I. Machavariani. 1982. *Sonantensystem und Ablaut in den Kartwelsprachen*. W. Boeder, tr. Tübingen: Günter Narr.

Gamqrelije, Tamaz. 1979. zmnis p'irianoba da valent'oba. [The morphological and syntactic valence of the verb]. In *saenatmecniero k'rebuli* [Studies in linguistics], ed. by Š. Jijiguri. Tbilisi: Mecniereba. (German translation in *Georgica* 4: 65-70.)

Gudjedjiani, Chato and Palmaitis, Letas. 1985. *Svan-English dictionary.* Delmar, NY: Caravan Books.

Harris, Alice. 1981. *Georgian syntax: a study in relational grammar.* New York: Cambridge University Press.

_____. 1984. On the origin of series markers in Kartvelian. *Folia Slavica* 7: 153-180.

_____. 1985.*Diachronic syntax: the Kartvelian case (Syntax and semantics 18).* New York: Academic Press.

Holisky, Dee Ann. 1981. *Aspect and Georgian medial verbs.* Delmar, NY: Caravan Press.

Hopper, Paul and Thompson, Sandra. 1980. Transitivity in grammar and discourse. *Language* 56: 251-299.

Imnaišvili, Ivane. 1957. *Saxelta bruneba da brunvata punckiebi jvel kartulši.* [Nominal declension and case function in Old Georgian]. *Jveli kartuli enis katedris šromebi* 4.

Jijiguri, Šota. 1940. K'oniunkt'ivis k'at'egoria rač'ulši ujvelesi kartuli mesame p'iris subiekt'uri supiksis p'roblemastan dak'avširebit. [The significance of the Rachan conjunctive for the problem of the earliest form of the third-person subject suffix in Georgian], *Ak'ad. N. Maris saxelobis enis, ist'oriisa da mat'erialuri k'ult'uris inst'it'ut'is moambe* V-VI: 103-107.

J̌orbenaje, Besarion. 1989. *Kartuli dialekt'ologia, I.* [Georgian dialectology]. Tbilisi: Mecniereba.

Kaǰaia, Lamara (ed.). 1984. *Xanmet'i t'ekst'ebi, 1: xanmet'i otxtavi.* [Khanmeti texts, 1: the Khanmeti gospels]. Tbilisi: Mecniereba.

Kaldani, Maksime. 1958. Svanuri enis balskvemouri k'ilos eceruli met'q'velebis taviseburebani [Characteristics of the Ecer subdialect of the Lower Bal dialect of Svan], *Iberiul-k'avk'asiuri enatmecniereba* 9-10: 201-218.

_____. 1978. Aorist'is c'armoeba svanurši [Formation of the aorist in Svan], *Iberiul-k'avk'asiuri enatmecniereba* 20: 150-161.

Kipshidze, Ioseb. 1914. Grammatika mingrelskago [iverskago] jazyka s xrestomateju i slovarem. *Materialy po jafetičeskomy jazykoznaniju* 8.

Klimov, G. A. 1964. *Ètimologičeskij slovar' kartvel'skix jazykov.* Moscow: Nauka.

_____. 1979. O donominativnom komponente kartvel'skoj jazykovoj struktury. In *Saenatmecniero k'rebuli* [Studies in linguistics], ed. by Š. Jijiguri, pp. 130-145. Tbilisi: Mecniereba.

Mač'avariani, Givi. 1974. Aspekt'is k'at'egoria kartvelur enebši [The category of aspect in the Kartvelian languages]. *Kartvelur enata st'rukt'uris sak'itxebi* 4: 118-141.

Melikišvili, Damana. 1977. Ink'luziv-eksk'luzivis k'at'egoriis gamoxat'vis ist'oriisatviskartul zmnaši [On the history of the expression of the inclusive/exclusive category in the Georgian verb]. *Macne* 4: 159-166.

Met'reveli, Teimuraz. 1978. Nochmals zur Kategorie von Inklusiv und Exklusiv im Altgeorgischen. *Georgica* 1: 23-29.

Mithun, Marianne. 1988. Lexical categories and the evolution of number marking. In M. Hammond & M. Noonan (eds.) *Theoretical morphology: approaches in modern linguistics*, pp. 211-234. New York: Academic Press.

Nichols, Johanna. 1985. The directionality of agreement. *Berkeley Linguistics Society* 11: 273-286.

Nozaje, L. 1974. Medioak'tiv zmnata c'armoebis zogi sak'itxi kartulši. [Some issues concerning the formation of medial verbs in Georgian]. *Iberiul-k'avk'asiuri enatmecniereba* 19: 25-53.

Oniani, Aleksandre. 1978. *Kartvelur enata ist'oriuli morpologiis sak'itxebi.* [Issues in Georgian historical morphology]. Tbilisi: Mecniereba.

_____. 1979. Mesame subiekt'uri p'iris morpema tanamedrove salit'erat'uro kartulši. [The third person subject morpheme in contemporary literary Georgian]. In *Saenatmecniero k'rebuli* [Studies in linguistics], ed. by Š. Jijiguri, pp. 153-161. Tbilisi: Mecniereba.

_____. 1984. Kartvelur enata diakroniuli morpologiis zogierti sak'itxis šesaxeb.[Concerning certain questions in Kartvelian diachronic morphology]. *Macne* 4: 192-202.

Sarǰvelaje, Zurab. 1984. *Kartuli salit'erat'uro enis ist'oriis šesavali.* [An introduction to the history of the Georgian literary language]. Tbilisi: Ganatleba.

Schmidt, Karl Horst. 1957. Eine südkaukasische Aktionsart? *Münchner Studien zur Sprachwissenschaft* 10. 9-24.

______. 1962. Studien zur Rekonstruktion des Lautstandes der südkaukasischen Grundsprache. *Abhandlung fur die Kunde des Morgenlandes* XXXIV #3.

______. 1969. Zur Tmesis in den Kartvelsprachen und ihren typologischen Parallelen in indogermanischen Sprachen. In *Saiubileo k'rebuli mijǧvnili Giorgi Axvledianis dabadebis 80-e c'listavisadmi*, pp. 96-105. Tbilisi: Tbilisi University Press.

______. 1978. On the reconstruction of Proto-Kartvelian. *Bedi Kartlisa* 36: 246-265.

______. 1982. Miscellanea Svanica. *Iberiul-k'avk'asiuri enatmecnierebis c'elic'deuli* (Annual of Ibero-Caucasian linguistics) 9: 62-73.

______. 1984. On aspect and tense in Old Georgian. *Folia Slavica* 7: 290-302.

Šanije, Ak'ak'i. 1953. *Kartuli gramat'ik'is sapujvlebi, 1: morpologia.* [Fundamentals of Georgian grammar, 1: Morphology]. Tbilisi: Tbilisi University Press.

Schanidse, Akaki (= Šanije, Ak'ak'i). 1982. *Grammatik der altgeorgischen Sprache.* H. Fähnrich, tr. Tbilisi: Tbilisi University Press.

Šarajenije, Tinatin. 1954. Saxelta mravalobiti ricxvis c'armoeba svanurši balszemouri k'ilos mixedvit [Formation of the plural number of the noun in the Upper Bal dialect of Svan]. *Iberiul-k'avk'asiuri enatmecniereba* 6: 189-203.

Talmy, Leonard. 1985. Lexicalization patterns: semantic structures in lexical form. In T. Shopen (ed.) *Language typology and syntactic description, Vol. 3: Grammatical categories and the lexicon.* Cambridge University Press.

Topuria, Varlam. 1967. *Svanuri ena, 1: zmna.* [The Svan language, 1: Verb]. Tbilisi: Mecniereba.

Tuite, Kevin. 1988. Number agreement and morphosyntactic orientation in the Kartvelian languages. Ph.D. Dissertation, University of Chicago.

______. 1989. The geography of Georgian *q'e*. In H. Aronson (ed.) *Non-Slavic languages of the USSR: linguistic studies*, pp. 283-302. Chicago: CLS.

______. 1990. Das Präfix *x-* im Frühgeorgischen. (To appear in *Georgica* 13).

Uhlenbeck, C. C. 1916. Het passieve karakter van het verbum transitivum in talen van Noord-Amerika. *Verslagen en Mededeelingen der Koninklijke Akademie van Wetenschappen. (Afd. Letterkunde).* 5, no. 2: 187-216.

Vogt, Hans. 1947. Le système des cas en géorgien ancien. *Norsk tidsskrift for sprogvidenskap* 14: 98-140.

Zorell, Franz. 1930. *Grammatik zur altgeorgischen Bibelübersetzung.* Rome: Pontificium Institum Biblicum

NOTES

[1] These designations, borrowed from Šanije [1953], indicate *form-based* verb-stem classes; the correlation with semantic characteristics is strong but by no means exceptionless. Class A includes transitives and medial (medioactive) verbs (more about them below). Class P is subdivided into four formal subclasses, and includes inchoatives, statives, passives and many verbs of motion.

[2] The reader will note that this is a highly tenuous proposal, since no traces of an inclusive/exclusive opposition in Set S, or of a cognate for the Svan S_{1incl} prefix *l*- have been attested in the other Kartvelian languages. It should also be taken into account that Svan has preserved a sizeable number of archaic features, compared to the other members of the family [Schmidt 1978].

[3] Kaldani [1978] has proposed, on the basis of umlaut patterns in the Svan aorist conjugation, that in Prehistoric Svan, a suffix *-*a* , cognate with the Georgian-Zan Group C S_{3sg} suffix, was added to the 3sg and all plural forms in the aorist. Concerning the Svan conjunctive S_{3sg} suffix -*s* , see Schmidt [1982:70].

[4] According to a widely-held hypothesis first proposed by Čikobava [1940], the Common Kartvelian S_3 marker was a *prefix*, giving prefixal agreement for all three persons in both Set S and Set O (see also Topuria [1967:8]; Andğuladze [1968:178-200]; Oniani [1984:167-85]). The only Kartvelian language where an S_3 prefix occurs is Svan. Topuria [1967:2-3] lists four verbs which take the prefix *l*-, one of which is the copular verb shown above. In addition, a large number of verbs in certain Svan dialects contain the elements -*y*- and -*n*- (phonological variants of *l*-) in the Set S 3rd person [Kaldani 1958 §14]. Čikobava [1940:41] argued that Svan S_3 *l*- and its phonological variants are reflexes of the Common Kartvelian S_3 prefix **n*- . This prefix is also hypothesized to be the ancestor of the Georgian-Zan S_3 marker -*n*, having undergone a shift from prefixal to suffixal position in Common Georgian-Zan. This suffix originally marked person only and not number, like the prefix from which it was descended. Oniani [1978:185] believes that the S_3 prefix was still used in Common Georgian-Zan when the realignment of tense-modal suffixes as person markers took place. The prefix, now redundant, disappeared soon afterwards. Recently Schmidt [1982] challenged this hypothesis. In his view, agreement with 3rd person arguments was not marked on the Common Kartvelian verb at all, as is the case in some modern Northwest Caucasian languages [1982:65-7]. The development of S_3 marking took place later, and followed different pathways in Prehistoric Svan and Common Georgian-Zan. In the former, an *l*-prefix related to the preverb *la*- was reinterpreted as a person marker [1982:68-9], and in the latter tense-modal suffixes were reinterpreted for the same purpose.

[5] Certain instances of number-agreement neutralization might be interpreted as traces of this earlier function of the S_3 suffixes. In 9th-13th century Georgian texts the Group D suffix -*ed* occasionally appears as a 3sg as well as 3pl Set S marker in certain screeves [Arabuli 1984:52-61; Sarǰvelaje 1984:394-9]. In the unusual

Mountain Rachan subdialect spoken in the village Glola, the S_3 suffix used in the conjunctive and optative screeves *(-n)* does not code number: *man//mat kn-a-n̲* "s/he:ERG//they:ERG should do it" [Jijiguri 1940].

[6] I am grateful to the anonymous reviewer of this article for emphasizing this point.

[7] In view of the phonological identity of the S_2 and O_3 prefixes, it cannot in fact be decided on formal grounds which prefix appears in this verb.

[8] In the early Old Georgian (Khanmeti and Haemeti) texts the O_3 marker preceded the S_1 marker: *x-w, h-w*. In later Georgian the S_1 morpheme always comes first: *v-h-* , *v-s-*.

[9] "Die Ableitung von der 3.Sg. als Basisform erfolgte dann in Protokartvelischen für die ersten beiden Personen des Singulars unabhängig von den drei Personen des Plurals" [1982:70]. Note that this reconstruction is based solely on Svan evidence (see note 2).

[10] If something like the 'reshuffling' hypothesis illustrated in Table 5 corresponds to reality, it is likely that stem$_0$ contained the tense-mood morphemes which later evolved into the Georgian-Zan S_3 suffixes.

[11] In the composition of the verb in the modern Kartvelian languages, one or more directional prefixes (preverbs) precede the person marker. These, together with the verb stem and certain other derivational morphemes, compose the lexical listing of the verb (semantically, Kartvelian preverbs resemble the adverbial particles in English verb-particle constructions). Evidence from Old Georgian and Svan indicates that at one time these preverbs were distinct sentential constituents [Schmidt 1969].

[12] The underlying form of the plural agreement morpheme is *en* according to Šanije [1982:75], with syncope of the /e/ in specified environments [Harris 1985:171,210]. It is important to keep in mind that the *-en-* morpheme to be discussed in this and the following sections is not claimed to be cognate with any of the Georgian-Zan S_3 suffixes containing /n/, with the partial exception of the Zan *-n-* element analyzed in §2.3.(B^2) below.

[13] Among the perfect-series screeves, the present perfect is the one exception: it never manifests *-en-* agreement. The present perfect is also formed from a different stem (the stative-passive present stem) from that used for the other perfect-series screeves (which employ the *e*-prefixed passive aorist stem) [see Harris 1985: 303].

[14] In this paper, the term 'absolutive NP' will be used to denote NPs functioning as subjects of intransitive verbs or direct objects of transitive verbs, regardless of their case marking.

[15] In addition to nouns with a segmentable plural suffix, the set of syntactically-plural NPs includes the 1pl and 2pl personal pronouns, and certain NP types which cannot undergo number declension: null anaphors and bare-stem

(c'rpelobiti) NOM-case NPs [Vogt 1947: 103; Sarǰvelaje 1984: 561-565; Harris 1985: 211-212]. In Old Georgian, intra-NP agreement and Set S agreement were also sensitive to syntactic, not semantic, plurality [Šanije 1982: 182ff; Tuite 1988 §§2.2.1, 3.2].

[16] Exceptions to this principle are attested in texts from the Old Georgian period [Sarǰvelaje 1984: 543ff], but the proportion of semantically — rather than purely formally — controlled number agreement remains small until the late Middle Georgian period [Tuite 1988 §3.4]. The period in the history of the Georgian literary language which saw the shift from *-n/-ta* to *-eb-* as the favored plural desinence likewise saw the decline of *-en-* agreement [Deeters 1930:65]. Those dialects which still make extensive use of the *-n/-ta* plural desinences (e.g., Pshavian, Xevsurian, Glola Rachan) are also the only dialects which preserve *-en-* agreement in Class A verbs.

[17] Nichols [1985: 274] defines agreement as "coincidence in grammatical categories, features, or feature values on two different words in a sentence, where one word has the category or feature for a principled reason and the other merely acquires it from the first."

[18] As such Georgian *da-* resembles the Polish preverb *po-* as described by Frajzyngier [1985: 97-8], which is correlated with both "plurality of action" (e.g., *po=stał w kolejce* "he stood in line <for quite some time>") and plurality of the local argument (e.g., *po=otwierał okna* "he opened <all> the windows." Also cf. *po=spadali* "they fell down" with the unacceptible **po=spadał* "he fell down").

[19] This verb is, from a purely formal point of view, a *double* causative, with the pluralizer intervening between the two causative formants. This semantically unmotivated doubling of causative formants is not rare in Svan [Topuria 1967: 234-235].

[20] E.g. Laz *ǧuru=n̲=an* "they are dying" (3sg *ǧuru=n)*. Čikobava [1936: 91; 1968: 163 n.3] interprets this pattern as the remnant of a more archaic, agglutinative mechanism of number agreement marking: 3pl = 3sg *ǧuru=n)* + plural *(-an)*. Contrast this with the pattern for Zan Class A verbs: *zumum=an* "they are measuring it" (3sg *zumum=s)*.

[21] "Im Rahmen der allgemeinen Grammatik greift der Gegensatz von Singularität und Pluralität über die Wortarten hinweg ... In der allgemeinen Grammatik wird man wohl nominalen Plural und Aktionsarten der verbalen Pluralität als kombinatorischen Varianten des Semems Pluralität bei Nomen und Verben bestimmen" [Dressler 1968: 94].

[22] Of the plethora of Svan nominal pluralizers inventoried, *-ar* is used by the largest class of nouns. Šarajenije [1954: 200] has hypothesized that the vowel quantity is a comparitively recent development, and that a form with a long vowel *(*-a:r)* is to be reconstructed.

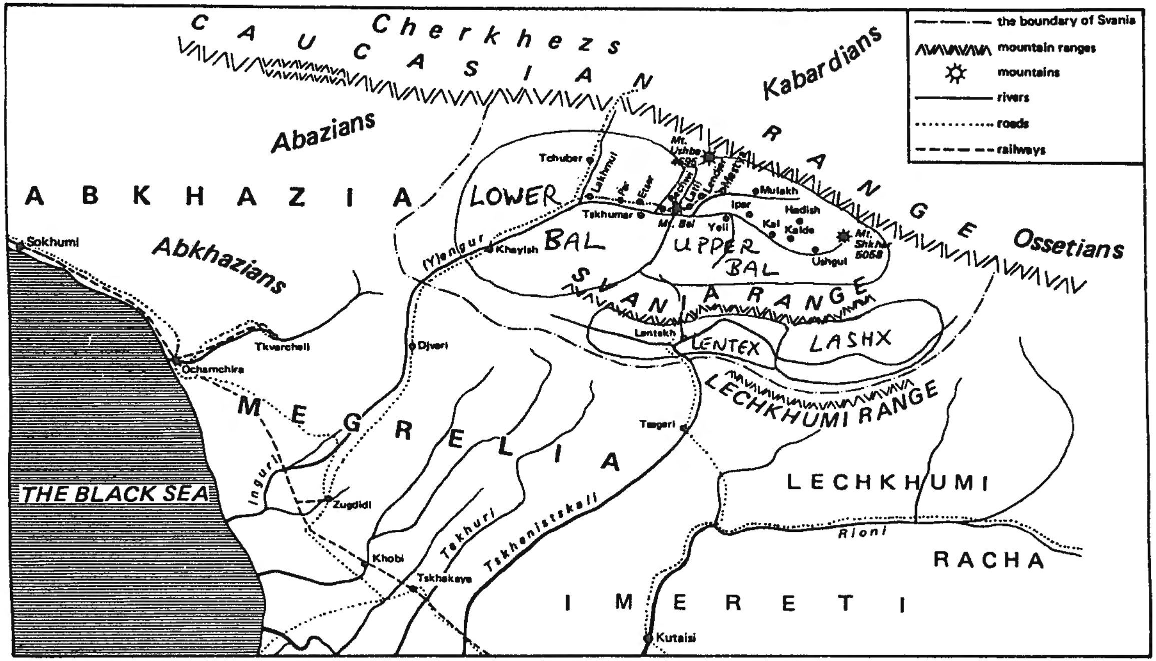
[23] The etymological connection between the Old Georgian pluralizers-*en*- (verbal) and-*n*- (nominal) is also accepted by Deeters [1930: 62] and Boeder [1979: 453].

[24] The exact function of -*eb*- in the attested Old Georgian period is difficult to establish, since it was used much less frequently than the -*n/ta* plural. Also, there are several cases of the-*eb*- and-*n/ta* plurals substituting for each other in different Old Georgian redactions of the same passage [Imnaishvili 1957: 293; Vogt 1947: 132ff].

[25] There is, in any event, strong evidence for a correspondence between the present-stem formants -*eb*- (Geo.) and -*ap*- / -*ep*- (Zan), on the one hand, and between the Georgian and Zan nominal pluralizers. In both cases the antecedent has been reconstructed as **eb* [Klimov 1964: 78]. This brings up the question of whether there were in fact several pluralizers common to nouns and verbs in Common Georgian-Zan, as there are in Svan.

[26] "The collectivizer **ev* was used in series I [present-series — KT] forms to mark durative aspect; it indicated that a set of actions was viewed as a whole, just as it indicated that a set of substantives was viewed as a whole" [Harris 1985: 207]. The semantic correspondence between durativity in verbs and collectivity in nouns was also noted by Dressler [1968: 75].

[27] It may be more faithful to the facts to consider the plural-iterative and collective-durative as two potentially crosscutting categories. This would allow, in theory, plurals of collectives, and iterative or distributive duratives (cf. Dressler [1968: 86-87]). Old Georgian nouns marked with both -*eb*- and -*n*- have been noted (e.g., *sopl-eb-n-i* 'village-PL_{eb}-PL_{n}-NOM' = 'several groups of villages'? [Imnaishvili 1957: 314; Sarǰvelaje 1984: 380]), and the Svan verb can have both a present-stem formant (also an old collective-durative?) and the distributive-iterative suffix *a:l*.



The Svan dialects [map reproduced from Ch. Gudjedjiani & M. Palmaitis *Upper Svan: Grammar and texts*. Vilnius: Mokslas, 1986; p. 15.]

INDO-EUROPEAN LENGTH AND BALTIC TONE

Steven Young
University of Maryland,
Baltimore Graduate School

It has been assumed since Saussure 1894 that original non-final long syllables (monophthongs and diphthongs) continue in Balto-Slavic under acute tone; circumflex is the tone of inherited normal diphthongs. The traditional handbooks such as Shevelov 1965 or Stang 1966 represent the acute : circumflex tonal opposition as arising phonemically with the shortening of long diphthongs and with the introduction of metatony (in the form of circumflex) on long monophthongs, but the phonetic details of the rise of Balto-Slavic tone are typically not considered; some accounts simply project the tonal features back to the proto-language.

The connection between Indo-European length and Balto-Slavic tone has in recent years been given a more productive interpretation by Kortlandt (1977), developing an earlier proposal of Vaillant (1936). In this approach, the broken-tone realization of acute found in certain East Baltic dialects is seen to continue a syllable containing at an earlier stage a laryngeal or glottal element (the latter as a result of Winter's law). But while providing a phonetic motivation for the origin of Baltic acute, this analysis represents a departure from Saussure, since acute here corresponds only to certain instances of Indo-European length; non-final longs which are not associated with Winter's law or laryngeal would be expected to carry the unmarked tone, circumflex, which arose in opposition to acute. There are, however, a number of forms containing long syllables conventionally described as instances of *Dehnstufe* which in fact appear with acute in Baltic (and Slavic; hereafter our primary concern will be Baltic; unless otherwise noted, all examples are from Lithuanian). Accordingly, we would need to posit a stage in the prosodic history of Baltic for which inherited length of whatever source is redundantly marked by a feature which continues in East Baltic as broken tone.

I. The first attempt to derive Baltic acute from a glottal feature is found in Vaillant (1936:114): "Les intonations balto-slaves supposent donc l'existence à date plus ancienne de deux sorts de longues et diphtongues: les unes douces [circumflex], à succession normale d'une montée et d'une descente de la voix; les autres rudes [acute], à coup de glotte ... " The glottal stop is identified with the laryngeal segments of Indo-European. More recently, Kortlandt (1977) has emphasized the connection between earlier laryngeal and the broken tone of Latvian (representing an earlier pretonic acute) and of Lithuanian Žemaitic dialects, corresponding historically to falling tone (tvirtapradė) of the standard language. In this revalorization of a segmental feature, laryngeal, to a prosodic feature, broken tone (containing a glottal stop), we have for the first time a phonetically motivated account of the origin of Baltic acute.

In addition to establishing a phonetic basis for Baltic tonogenesis, the approach of Vaillant and Kortlandt presents several other advantages:

A. Certain later prosodic phenomena of Lithuanian are more easily accounted for phonetically. For Saussure's law, for example, according to which ictus is attracted onto an adjacent acute syllable, explanations based on purely tonal or mora features lack phonetic plausibility. By positing a particular intensity (such as glottal stop) for acute syllables, the attraction of ictus is more readily accounted for.

B. Leskien's law is better motivated. According to Leskien's law, final acute monophthongs in Lithuanian were shortened, while final acute diphthongs became circumflex, two processes which are otherwise unrelated. If we assume loss of a glottal feature in all final acutes, however, we automatically get truncated vowels (note that acute syllables are typically shorter than circumflex syllables to begin with), as well as circumflex (unmarked) diphthongs. It is significant that those Žemaitic dialects which preserve broken tone acute lack reflexes of Leskien's law in diphthongs--final acutes here are preserved as such (with broken tone): *šokaû* 'jump (1sg.pres.)' (standard *šokaũ),* *nešê* 'take (2sg.past)' (standard *nešeĩ), darŷs* 'do (3fut.)' (standard *darỹs).*

C. It becomes easier to understand why Lithuanian borrowings from Slavic traditionally represent long vowels with circumflex rather than acute: acute, with its glottal feature, is marked; circumflex, without this feature, is unmarked.

D. Finally, the solution proposed by Vaillant and Kortlandt allows us to solve the riddle of the traditional handbooks, which present acute as falling tone in Lithuanian, but as rising tone in Latvian, Old Prussian, and Slavic (and vice-versa for circumflex). Traditionally, Lithuanian has been viewed as having undergone a complete reversal in tonal contours (Stang 1966:125). The mechanisms put forward to account for this process often seem rather contrived (for example Martinet, 1986:75ff., who assumes such a change for Latvian).

Vaillant has suggested that the tonal contour associated with Baltic acute is in fact secondary, connected with the location of the glottal feature within the syllable. He assumes a syllable-final position, occasioning an "élévation brusque de l'intensité et du ton" (1950:245). In the Žemaitic broken tone syllable (representing the strongest dialectal exponent of acute) the glottal interruption comes early; such syllables are characterized by a sharp rise in pitch to the glottal feature, followed by a slower fall, yielding an overall falling contour. Proceeding eastward and southward across the Lithuanian dialect territory, there is a gradual diminution in the salience of acute: the glottal feature is lost, leaving the concomitant tonal contour, which, in turn, is gradually levelled, until the situation of the eastern dialects (as well as the urban colloquial language) is reached, where the tonal opposition is neutralized on monophthongs.

The falling tone considered normative for the standard language represents the situation in the southwest Aukštaitic dialect, on which the standard is based. The long falling tone of the handbooks, then, is simply one of the geographically intermediate stages in the Lithuanian realization of acute, the extremes being glottal stop in the West and neutralization of tone in the East. Circumflex has developed as non-acute, originally non-broken, and varies dialectally in its tonal contour. In dialects with broken tone acute, the circumflex is rising-falling, and in dialects with falling acute, it is level to rising, the latter especially for diphthongs. The comparativists' selection of falling tone to represent original Lithuanian acute and rising tone to represent original circumflex is an error in method, influenced by the prestige of one dialect over others. What should be set up on the Lithuanian side of the equation is not falling tone for acute, but broken tone, and for circumflex, non-broken tone, with its various dialectal realizations.

It is likely that acute was originally a broken tone in Latvian as well. Significantly, the broken tone in central Latvian dialects represents historically pretonic acute (note that various Žemaitic dialects also have pre- and post-tonic broken-tone acutes). At present, it is not clear why in original stressed acute syllables in Latvian the glottal feature would disappear, but it seems that the original concomitant pitch contour still remains, perhaps somewhat adapted, as rising tone. The true opposition is thus glottal feature for both Latvian and Lithuanian, with subsequent separate developments of tonal contours not only within each of these languages, but across their respective dialects as well.

II. Although the majority of instances of acute syllables in Baltic are associated with an earlier laryngeal, there still remain cases of what Stang (1970:216) has termed "theoretical metatony," where acute, rather than expected circumflex, occurs in a non-laryngeal base. Most of these exceptional cases, such as Balto-Slavic **sḗd-* (OInd. *sad-*) 'sit,' **bḗg-* (Gk. *p*^{h}*ébomai*) 'run,' **éd-* (OInd. *admi*) 'eat,' have been shown by Winter (1976) to reflect an earlier distribution of stem-final voiced stops: a syllabic followed by an original voiced unaspirated stop is lengthened (with concomitant acute) in Balto-Slavic; in other cases, the original quantity is retained.

Kortlandt (1976) has brought Winter's observation into the broader framework of Baltic tonogenesis by connecting it with the postulation of a series of glottalic stops (in Winter's formulation, the traditional voiced stops). The glottalic feature characterizing these stops eventually merged in Balto-Slavic with the reflex of the inherited laryngeals (Kortlandt 1977:322) as acute tone. In this way the acute of the following bases also receives a phonetic motivation: **u̯eid-* 'to see': *véidas* 'face'; **melǵ-* 'to milk': *mélžiu* 'milk (1sg.pres.)'; **s(t)erg-* 'to watch over' : *sérgiu* 'guard (1sg.pres.)'; **skeid-* 'to split, separate': *skéidžiu* 'dilute (1sg.pres.)'; **i̯eug-* 'to join': *jùngiu* (with intrusive -*n*-) 'join (1sg.pres.)'; **u̯ed-* 'wet; water': *vánden-* (with intrusive -*n*-) 'water'; **leig-* 'to bind': *láigonas* 'brother-in-law (wife's brother)'; **perd-* 'to fart': *pérdžiu* 'fart (1sg.pres.)' (cf. Gk. *pérdomai*).

III. As we have seen, the majority of Baltic acute syllables can be given a purely phonetic account in terms of laryngeal theory or

Winter's law. There still remains, however, a residue of inherited forms with acute tone for which no phonetic explanation is possible, and on these grounds Stang (1952:449) finds it difficult to accept Vaillant's theory as it is stated. Among such forms are those with Dehnstufe length, originally representing compensatory lengthening due to the loss of a segment in monosyllabic word-forms of stems in nasal, liquid, or *-s-* (the nom.sg. of consonant stems, for example; see Szemerényi 1970:109 ff.).

Where inherited Dehnstufe length appears in final (desinential) position, the Baltic prosodic reflex is circumflex tone, as for length by contraction (e.g. gen.sg. ā-stem *galv-õs* 'head' < **-ā̃s* < **-eH*$_2$*os*), or in normal diphthongs (gen.sg. *u*-stem *sūn-aũs* 'son' < **-ous*). Instances where an earlier stem-final -V̄R- (more rarely -VC-) sequence is reinterpreted as a marker of nominative case include: *akmuõ* 'stone' (**akmōn*), š*uõ* 'dog' (**ḱuōn*), *sesuõ* 'sister' (**sesōr*), *mótė* 'mother' (**mātēr*), *ménuo* 'month' (**meHnōt* [cf. Specht 1944:344]; circumflex tone on the unstressed syllable in *mótė* and *ménuo* is indicated by a failure of Leskien's law to apply).

When, on the other hand, these long syllables appear within a stem, the Baltic (Balto-Slavic) reflex appears to be acute; thus where an original final long has become word-internal by virtue of a shift from the consonantal stem class to other stem classes: *i*-stems *žvér-į* 'beast' (Gk. *t*^{h}*ēr*, with lengthened grade generalized throughout the stem [see Dybo 1981:40]), *lū́šis* (C-stem remnant in gen. pl. *lūš-ų*) 'lynx' *(:*leuḱ-),* Slvc. *myšĭ* 'mouse' (**mūs* [Szemerényi 110]). In *geluõ* (=*geluonìs*) 'sting,' with final circumflex, but Latv. *dzeluônis* 'id.' (representing an extension of old *-n* stem; see Endzelin 1922:241) with stem-final acute, we have an example in which -V̄̃R- appears both internally as well as in word-final position. Indeed, Arumaa 1985:23-4 sees the acute suffixes *-úonis* and *-énas* as thematic continuations of earlier final *-ōn*, *-ēn*, with paradigmatic generalization of Dehnstufe length.

Phonetically unmotivated acute length is also found in the first syllable of the isolated forms *úolektis* 'ell' (**ōlekt*; compare Latin *ulna* < **olenā* for short vowel [Watkins 1965:121]); *úosis* 'ash[-tree],' Slvc *(j)ăsen- (compare Latin *ornus* < **osi-no-* for short vowel [Watkins 1965:121]); in derivatives such as *akìs* 'eye' : *akýtas* 'porous' (*-i-* : *-ī-*; cf. Specht 1944:326), *gaũras* 'hair' : *gaurúotas* 'shaggy' (*-o-* : *-ō-*), if in fact these do not represent sequences with laryngeal (alongside the *gaurúotas* type, Watkins 1965:120 considers infinitives in *-úoti* among *o*-stem derivatives with secondary lengthening); and, on the same principle, but undoubtedly as a secondary development, in compound derivations of the type *pėlẽ* 'mouse' : *pėlékautai* 'mouse-trap,' *žasų̃* 'goose (gen.pl.)' : place-name *žasū́gala*, **naviẽs* 'death (gen.sg.),' cf. La *nâve* 'id.' : *navíeskaulis* 'tumor' (Skardžius 1941:406). Also here are the inherited *TáRTā* formations of the type *várna* 'crow' (Mathiassen 1974:208, 221), *vìlkė* 'she-wolf' (:*vĩlkas* 'wolf'; Stang 1966:158).

Finally, the Latvian gerund, historically the dative singular of a present participle, shows acute tone on the formant *-uõt* (Endzelin 1922:28); this may well represent an instance of inherited Dehnstufe length (**-onts* > *-ōn*, with *-t-* reintroduced in Latvian) preserved internally: *vagājuõt* 'furrowing,' *velkuõt* 'pulling,' *liekuõt* 'putting'; the (broken-tone) acute in the second syllable of *tũkstuôtis* 'thousand' is similarly accounted for.

There is, it should be noted, a systematic set of exceptions to the above argument: as Kortlandt 1985 has shown, a number of rebuilt root nouns (typically verbal in nature), in which the Dehnstufe is apparently inherited from the proto-language, show circumflex rather than acute tone in the base (see also Arumaa 1985:50). In the present framework I have no explanation for these stems.

• • •

In order to account for the residual acutes presented above without abandoning the richly explanatory interpretation of Baltic acute as the reflex of an earlier glottalic feature, a type of analogical extension must be invoked. At a stage of early Baltic (Balto-Slavic) preceding the contraction of final disyllabic sequences such as *-eH_2os* to monophthongs (later characterized by circumflex), the glottalic feature resulting from laryngeal and Winter's law was extended to all long syllabics (both internal and word-final). Subsequently, this feature was lost before a final resonant (such a rule is already required to account for the lack of acute in the accusative singular of ā-stems: *rañką* 'hand,' without Saussure's law); finally, a word-final resonant was lost after a long vowel: PIE **mātēr* > **mâtêr* > **mâtēr* > Lith. *mótė*. The development presented in this way accounts for the different tonal treatment of Dehnstufe length internally and word-finally.

The analogical extension of a glottal feature beyond its original domain to all vowels of a certain type is not without precedent in the historical prosody of Baltic. Just such a process must be posited for Lithuanian Žemaitic dialects which extend broken tone to "secondary acute" diphthongs formed from original disyllabic sequences containing a resonant. As the second, unstressed, syllable in such sequences is lost, a stressed and unstressed mora become tautosyllabic (CVR̭VC > CVR̥C), with a resulting falling tonal contour, in examples such as *gêr̥s* 'good' < *gẽr̭as*, *mâl̥* 'grind (3pres.)' < *mãḽa* (Grinaveckis 1973:89). The concomitant tonal contour (falling) in Žemaitic acute syllables serves as a point of departure for the extension of the broken tone acute to such newly closed syllables. By thus appealing to a model already present in East Baltic, we are able to uphold Saussure 1894 and satisfy Stang's reservations concerning Vaillant's theory of acute tone as ultimately glottal (Stang 1952:449). We thereby preserve a phonetic motivation for Baltic acute, an approach which Stang himself had termed "séduisante."

REFERENCES

Arumaa, Peeter. 1985. Urslavische Grammatik. Vol. III: Formenlehre. Heidelberg: Carl Winter.

Dybo, Vladimir Antonovič. 1981. Slavjanskaja akcentologija. Moscow: Nauka.

Endzelin, Jan. 1922. Lettische Grammatik. Riga: A. Gulbis.

Grinaveckis, Vladas. 1973. Žemaičių tarmių istorija (fonetika) [A History of the Žemaitic Dialects]. Vilnius: Mintis.

Kortlandt, Frederik. 1976. Comment on W. Winter's Paper. Recent Developments in Historical Phonology, ed. Jacek Fisiak, 447. The Hague: Mouton.

_____. 1977. Historical Laws of Baltic Accentuation. Baltistica 13(2).319-329.

_____. 1985. Long Vowels in Balto-Slavic. Baltistica 21(2).112-124.

Martinet, André. 1986. Des steppes aux océans. L'indo-européen et les Indo-Européens. Paris: Payot.

Mathiassen, Terje. 1974. Studien zum slavischen und indo-europäischen Langvokalismus. Oslo: Universitetsforlaget.

Mayrhofer, Manfred. 1986. Indogermanische Grammatik. Vol. 1, part 2. Heidelberg: Carl Winter.

Saussure, Ferdinand de. 1894. À propos de l'accentuation lituannienne. Mémoires de la Société de linguistique de Paris 8, 425-446.

Shevelov, George. 1965. A Prehistory of Slavic. New York: Columbia University.

Skardžius, Pranas. 1941. Lietuvių kalbos žodžių daryba [Lithuanian Word Formation]. Vilnius: Lietuvių Mokslų akademija.

Specht, Franz. 1944. Der Ursprung der indogermanischen Deklination. Göttingen: Vandenhoeck & Ruprecht.

Stang, C. 1952. Review of Grammaire comparée des langues slaves, by André Vaillant. Norsk Tidskrift for Sprogvidenskap 16.433-451.

_____. 1966. Vergleichende Grammatik der baltischen Sprachen. Oslo: Universitetsforlaget.

_____. 1970. Opuscula linguistica. Oslo: Universitetsforlaget.

Szemerényi, Oswald. 1980. Einführung in die vergleichende Sprachwissenschaft. Darmstadt: Wissenschaftliche Buchgesellschaft.

Vaillant, André. 1936. Le problème des intonations balto-slaves. Bulletin de la Société de linguistique de Paris 37.109-115.

_____. 1950. Grammaire comparée des langues slaves. Vol. 1. Lyon-Paris: IAC.

Watkins, Calvert. 1965. Evidence in Balto-Slavic. Evidence for Laryngeals, ed. Werner Winter, 116-122. The Hague: Mouton.

Winter, Werner. 1976. The Distribution of Short and Long Vowels in Stems of the type *èsti* : *vèsti* : *mèsti* and OCS *jasti* : *vesti* : *mesti* in Baltic and Slavic Languages. Recent Developments in Historical Phonology, ed. Jacek Fisiak, 431-446. The Hague: Mouton.